AN INTRODUCTION TO STATISTICAL METHODS

THE MACMILLAN COMPANY
NEW YORK · BOSTON · CHICAGO · DALLAS
ATLANTA · SAN FRANCISCO

MACMILLAN & CO., LIMITED
LONDON · BOMBAY · CALCUTTA
MELBOURNE

THE MACMILLAN CO. OF CANADA, LTD.
TORONTO

AN INTRODUCTION

TO

STATISTICAL METHODS

A TEXTBOOK FOR COLLEGE STUDENTS
A MANUAL FOR STATISTICIANS
AND BUSINESS EXECUTIVES

BY

HORACE SECRIST, PH.D.
ASSOCIATE PROFESSOR OF ECONOMICS AND STATISTICS
NORTHWESTERN UNIVERSITY

New York
THE MACMILLAN COMPANY
1917

Set up and electrotyped. Published December, 1917.

Norwood Press
J. S. Cushing Co. — Berwick & Smith Co.
Norwood, Mass., U.S.A.

To

THE MEMORY OF

MY FATHER

JACOB M. SECRIST

PREFACE

The following chapters are an attempt to work out an introductory, but at the same time a comprehensive, text on statistical methods for the use of college students and students in colleges of business administration. They are also intended to supply the need for a fundamental treatment of the methods of statistical investigation and interpretation. Statistical methods are regarded as means rather than as ends, as constituting simply one phase of general methodology, and as including not only methods of analyzing but also of collecting and assembling statistical data. The methods discussed are of general application although the illustrations, for the most part, are drawn from economic and business fields.

The order of treatment is the same as that followed in the planning and analysis of a statistical problem, and it is hoped that statisticians, business executives, and students of statistical methods generally will find the volume not only a compendium of statistical procedure but also a guide in the process of logical statistical analysis. Emphasis is given to the necessity of a clear formulation of the problem in mind, to the meaning, collecting, and assembling of data, and to the necessity of a rigid interpretation and use of units of measurements. All of these steps are held to be preliminary but indispensable to the formulation of a statistical judgment, and to the employment of the refinements of mathematical analysis which alone are too generally associated with "statistical methods."

The treatment is non-mathematical for several reasons, chief of which are, that the mathematical phases of the subject are treated in other places, and that there seems to be an urgent need for a fundamental discussion of the non-mathematical, but not less vital, processes in statistical investigation and analysis. Experience in teaching statistics both to college students and business men, as well as in conducting statistical investigations, has demonstrated the need for such a treatment. It has been the aim at every stage of the discussion to develop the "why" of statistics, and concretely to relate methods to the problems of public and private economics.

The bibliographical aids at the close of the several chapters are not meant to be inclusive, but are chosen because of their value to students and others as collateral reading. A discussion of certain of them along with the text treatment, and in the light of the laboratory problems assigned, has proved helpful in the author's classes.

I am indebted to Professor Willard E. Hotchkiss, formerly Dean of the Northwestern University School of Commerce, and to Professor John F. Hayford, Dean of the Northwestern University College of Engineering for reading parts of the manuscript and for offering many helpful suggestions for its improvement. Most of all I am indebted to my wife who has materially lightened the burden of proof-reading, and who, at all stages in the preparation of the volume, has been a constant source of encouragement.

HORACE SECRIST.

NORTHWESTERN UNIVERSITY,
EVANSTON, ILLINOIS,
November, 1917.

CONTENTS

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

CHAPTER V

CHAPTER VI

CHAPTER VII

CHAPTER VIII

CHAPTER IX

CHAPTER X

CHAPTER XI

CHAPTER XII

LIST OF PLATES

AN INTRODUCTION TO STATISTICAL METHODS

CHAPTER I

THE MEANING AND APPLICATION OF STATISTICS AND STATISTICAL METHODS

I. Introduction

The necessity of basing economic and business judgments upon facts and of being able properly to collect and interpret them in connection with almost all of the different phases of economic activity is a sufficient general excuse for submitting a volume, the main purpose of which is a study of the principles governing the collection, analysis, and synthetic treatment of numerical data. More and more economic and business policies are being advocated after careful study of facts, and those affected by these policies are more and more frequently asking that they be given these same facts in a definite and understandable form. The tendency to base a case, to advocate a far-reaching change, to stand sponsor for a program or to agitate a reform, upon an appeal to natural rights, or to the innate goodness or perversity of human nature, is rapidly being overcome. Appeal to the force of custom and tradition alone no longer suffices as a basis for an economic program. If considered at all it is only to explain or appraise the facts involved. What is now being done is more closely to observe the reaction of forces under given

conditions, to enumerate the frequencies with which each reaction occurs, to test the closeness with which a given result follows a given cause, and to allocate and associate causes and effects generally.

Economic life and business dealings are more and more being determined by precise findings, while governmental policies are coming to be supported or condemned by an appeal, not alone to custom, but to their respective benevolent or malevolent effects. Business ventures are being pursued on narrow margins of profit and the effects of a policy determined by elaborate analysis of the results properly attributable to it. Ours is the age of the concrete and the realistic, as contrasted with the abstract and the metaphysical. We are no longer content to conjure up an "economic man" and to postulate his reactions under all circumstances. Explanations for economic and social phenomena, as the existence of a wage class, strikes, lockouts, unemployment, industrial disease, industrial accidents, premature death, panics, economic wastes, business failures, etc., are no longer sought in the wrath of God, in the movements of heavenly bodies, in the wickedness and perverseness of a people, in the sacredness of natural rights, nor looked upon as the necessary and unavoidable consequence of the present scheme of production and distribution. These phenomena, we have come to see, have their explanation in economic and social practices and usages, and we are able to determine their causes and effects, as well as to suggest methods of changing them or avoiding their consequences by a study of facts. Many of these may be expressed numerically and studied statistically.

Our study is primarily one of methods — methods in the collection and utilization of numerical data to throw light upon economic and business problems. It attempts to reduce to a workable basis the principles of statistical analysis

and to illustrate their force and the methods of applying them to concrete problems. The needs and problems of the student and of the man of affairs who is placed in a position of responsibility, where the exercise of judgment growing out of business experience is necessary, have been kept constantly in view. It is assumed that the man of affairs desires to act rationally upon the basis of facts at his command or capable of being acquired which bear upon his problems, and to formulate his judgments in their light and in a scientific manner. It is also assumed that the student desires to get at the foundation of his problem, to understand it in all its bearings, to be able to marshal all the facts which apply to it and to appraise their worth. But it is acknowledged that the statistical is only one approach to the understanding of a problem, and it is one of the main purposes of what follows to establish it in its proper position. Too much faith is often placed in the efficacy of statistics to "prove things." Reasoning from other angles than the statistical is too frequently dispensed with — if not utterly ignored — on the part of the uninformed when "statistics" can be utilized, notwithstanding the fact that the "statistics" may have no application, may be incomplete, unrepresentative, and questionable in origin, and that the problem cannot be understood by an appeal to its numerical side. Loose reasoning and hasty judgments are even less defensible when statistics are appealed to to support a contention than when they are ignored, for the reason that they seem to carry a finality and to suggest a nicety of conclusion not generally associated with a less precise method of approach.

"A given economic fact is the result of numerous complex forces, many of which are in a state of constant variation and react upon one another; and of these forces only a few can be adequately de-

scribed by the method of statistics. Consequently these few are often quoted as if they were the only active causes whereas the effect attributed to them is probable only on the assumption that all other causes remain unchanged or suspended. . . . Statistics, even when compiled accurately, though often absolutely necessary for a complete solution of a problem, do not in themselves provide that solution, but are to be used in conjunction with evidences of other kinds."[1]

Ignoring this fact, fallacies both of observation and inference[2] abound, and it is these to which the following discussion is addressed.

Newsholme, summarizing Quételet, lays down four rules for statistical studies:

"Never have preconceived ideas as to what the figures are to prove.

"Never reject a number that seems contrary to what you might expect, merely because it departs a good deal from the apparent average.

"Be careful to weigh and record *all* the possible causes of an event, and do not attribute to one what is really the result of a combination of several.

"Never compare data which have nothing in common."[3]

Without attempting at this time in any complete manner to formulate rules for statistical studies, the point of view upon which the treatment proceeds may be clearly indicated by calling attention to certain well-marked tendencies among beginners in the use of statistics and statistical methods.

(1) The tendency to accept without serious question a plausible description of a given condition or state of affairs. *Ipse dixit* is often regarded as sufficient proof. The mere fact of data appearing in print, and particularly of their

[1] McIlraith, James W., *The Course of Prices in New Zealand*, 1911, p. 4 of *Introduction* by J. Hight.

[2] Newsholme, Arthur, *The Elements of Vital Statistics*, 3d Ed., p. 294.

[3] *Ibid.*, pp. 292–293.

being in tabulated form — the finality of a statistical table is often magical — is frequently sufficient to insure their value and to guarantee their application. Respect for age, for custom, or for a condition of status quo is really remarkable in the unsuspecting in spite of the "show me" attitude which seems to characterize our period.

(2) The tendency to employ data without knowledge of or regard for the units of measurements in which expressed, or their comparability or representativeness, and to draw conclusions from them which they were never intended to support. This is the tendency which has been popularly characterized as the ability to "prove anything by statistics." On the other hand, not infrequently a realization of the limits of the statistical approach serves to restrict the use of statistics in cases where in reality the method is defensible. In such cases ignorance or distrust makes impossible the use of a valid instrument of study.

(3) The tendency to disregard detail, — or to regard it as "detail" which somehow will take care of itself and needs no especial attention, — to ignore statistical cautions respecting the collection of data or the use of those already collected, — to speak in terms of statistical abbreviations, averages of all types, — to employ totals as if they were always more sacred and inviolate than the items which go to make them up, and to piece together statistical fragments, gleaned from widely different sources and compiled under widely different circumstances, into a beautiful mosaic which thoroughly proves or disproves a contention already held.[1]

[1] For an admirable discussion of the false uses to which statistical data will be put, even by those who are in a position to know their limits, when it is a question of making a case, see Bowley, A. L., "Statistical Methods and the Fiscal Controversy" in *The Economic Journal* (London), Vol. 13, 1903, pp. 303–313. In formulating the rules to be observed, Bowley says:

(4) Lack of ability definitely to formulate the purpose of a statistical study, to outline appropriate methods in order to serve the end desired, to define with precision the units employed in the measurements, and rigidly to limit the field to be covered, — in a word, lack of ability to plan and execute a statistical study.

(5) Lack of knowledge of the sources and value of secondary statistical material — material already collected, tabulated, summarized, and analyzed — and of primary statistical material — material in a crude, disorganized, undigested form available for collection and analysis.

(6) Lack of knowledge of the methods of statistical analysis and synthesis.

It is the primary purpose of this volume, together with readings and laboratory problems, to supply these deficiencies — to put the reader in possession of the information, tools, and skill whereby he can, in a measure, not only pass upon the merits of the statistical approach to economic and business problems, and appreciate the problems involved in statistical studies, but can also undertake them independently.

"Every statistical estimate should be considered in the light given by corresponding estimates for previous years.

"Every total should be homogeneous in that quality which concerns the argument.

"Where values are used, the effect of replacing them by quantities should be tested.

"The errors latent in the constituents which form an estimate should be examined, and their effect on the estimates should be tested with reference to the purpose for which the estimate is used. The maximum adverse errors should be calculated, to see if their concurrence would vitiate the result.

"The ideal measurement necessary to support each deduction should be conceived; and if the estimates accessible do not necessarily give the same view as the ideal measurement, they should be rejected.

"When the sufficiency of statistics as estimates is established, the arguments based on them should be bound to the statistical results by the ordinary rules of logic." *Ibid.*, p. 312.

II. The Meaning and Application of Statistics and Statistical Methods

1. *The Meaning of Statistics and Statistical Methods*

Statistics is generally thought of from two points of view: first, as series of isolated numerical facts; and second, as methods involving the collecting, sorting, classifying, tabulating, summating, and comparing enumerated facts for the purpose of describing or explaining phenomena with which enumerations deal. Viewed solely in the first light, statistics is little more than arithmetic, and as such has little or no interest for us. From the second point of view, statistics closely approaches logic, concerned as it is with the processes and methods of formulating and testing conclusions from premises resting solely upon numerical bases.

Obviously, however, the function, process, or method side, *i.e.* the application of methods of analysis in order to suggest the inferences and conclusions to be drawn — cannot be divorced from the enumeration side, since it is the latter which helps to shape the premise the consequence of which it is desired to formulate. The conditions governing enumeration, such as the units and accuracy of measurements or enumeration, the completeness or representative character of the samples, etc., are vital and largely determine the methods to be employed in analysis. The adequacy of a tool, or the perfection of a machine, to speak analogously, is quite as important in the determination of a product as is the method of its utilization. However, skillful use may partly compensate for a poor tool, as skillful discrimination in statistical analysis may tend to counteract the error following from crude or defective enumeration. Statistics, as method, is as vitally concerned with enumeration as with

the process and manner of analysis and synthesis, and in what follows the principles of methodology are extended to both phases of statistical study.

In definitions of statistics the emphasis has been variously placed. Bowley has called it the "science of averages" [1] as well as "the science of counting." [2] The first definition emphasizes one device for statistical abbreviation; the other calls attention to the enumeration which precedes analysis. In another place, Bowley defines *statistics* as "numerical statements of facts in any department of inquiry, placed in relation to each other," and *statistical methods* as "devices for abbreviating and classifying the statements and making clear the relations." [3] Yule defines *statistics* as "quantitative data affected to a marked extent by a multiplicity of causes," and *statistical methods* as "methods specially adapted to the elucidation of quantitative data affected by a multiplicity of causes." [4] Still others, using the terms with less precision, and in a less scientific sense, have sought to identify statistics with graphic methods — to convert the science into an art. With the latter purpose we have little sympathy, yet due attention is later given to graphic methods as a means of statistical presentation.

We shall use the term *statistics* as meaning *aggregates of facts, "affected to a marked extent by a multiplicity of causes," numerically stated, enumerated, or estimated according to reasonable standards of accuracy, collected in a systematic manner for a predetermined purpose, and placed in relation to each other.*

This definition seeks to emphasize the fact that before

[1] Bowley, A. L., *Elements of Statistics*, p. 7.
[2] *Ibid.*, p. 3.
[3] Bowley, A. L., *Elementary Manual of Statistics*, p. 1.
[4] Yule, G. U., *An Introduction to the Theory of Statistics*, p. 5.

numerical data can be termed "statistics" they must bear evidence of having been collected in accordance with at least the rudiments of scientific method and for a definite purpose. It is necessary to insist that these conditions be fulfilled in order to know anything about the units of measurements employed and the scope and representativeness of the facts given numerical expression. Data not fulfilling these conditions may be numerical but they are not statistical. Too often "statistics" degenerate into "figures," and so-called "statistical bureaus" into nothing but "figure factories." Moreover, as Yule points out, "the term statistics is not usually applied to data, like those of the physicist, which are affected only by a relatively small residuum of disturbing causes." [1] Hence our reason for insisting, with Yule, upon the last-named condition. The requirement that statistics should conform to systematic and scientific methods of enumeration or estimation seems to connote the further condition that numerical facts are statistics only when "placed in relation to each other." [2] Stray and loose bits of information, gleaned here and there from indiscriminate sources, hearsay and unrelated material, while numerical in character, can be termed statistical only by a confused and unscientific use of terms. If they are capable of verification, if they take on homogeneity and assume regularity, then they may properly be classified as statistics.

The expression *statistical methods is used to include all those devices of analysis and synthesis by means of which statistics are scientifically collected and used to explain or describe phenomena either in their individual or related capacities.*

[1] *Ibid.*

[2] Bowley, A. L., *Elementary Manual of Statistics*, p. 1.

2. *The Application of Statistical Methods*

Statistics may be collected on most topics, but the employment of statistical methods in their study is not of universal nor of equal validity. At best the statistical is but *one* of many approaches in the explanation of phenomena. Its limitations are definite and certain, and its use in *all* cases should in no sense be considered valid. Statistics may often be used to corroborate conclusions arrived at by other methods, and it is in this respect, probably, that their greatest value lies. Many questions do not admit of statistical treatment at all; while respecting others, statistical considerations are of minor or of no consequence. The limitations of such methods are appreciated and clearly determined only after considerable experience, and it is one of the purposes of the volume to supply this training and experience.

But this does not mean that their function is narrow and restricted. Both inside and outside of business, occasions are daily arising where statistical facts are indispensable as bases for decisions of policy, methods, etc. By means of them improvident and unbusiness-like methods may be detected, and new policies, savings, and projects suggested. The importance now assigned to proper methods of accounting and cost keeping in business is proof that this fact is being realized, and that definite knowledge of costs, profits, expenses, etc., is necessary to success. Accounting is concerned with the value aspect of these problems; statistics relates to the numerical or quantitative aspect whether value or some other unit is chosen as a measure of activity. These means of scientifically analyzing business are complementary. The need to-day is an appreciation of facts, an ability to observe the conditions which produce them, and a determination logically and scientifically to piece them together in

such a way that they will serve as rules of business guidance. The problem, therefore, involves the establishment of units of measurements, analysis of activities according to these units, and the formulation of policies on the basis of the observations.

The application of statistics and statistical methods to economic and to business problems is sufficiently emphasized, at this place, by merely calling attention to a few of the various fields in which they may be employed. The discussion, illustrations, and problems subsequently introduced serve definitely to bring out the detailed application.

(1) Application within Business Units.

- *a.* Analysis of sales and sales possibilities by districts, by periods, by products, etc.
- *b.* Analysis of production by departments, processes, etc.
- *c.* Analysis of employment as to rapidity of turnover, scale of payment, labor supply, welfare work, etc.
- *d.* Analysis of production and factory organization.
- *e.* etc.

(2) Application without and between Business Units. Affecting,

- *a.* Consumption
  - (*a*) family budgets.
  - (*b*) price phenomena.
  - (*c*) etc.
- *b.* Production
  - (*a*) capital and labor employed, the absolute amounts and proportions.
  - (*b*) expenses incurred and their distribution.
  - (*c*) materials used — amounts and values, and their distribution.
  - (*d*) products created — amounts and values, and their distribution.
  - (*e*) etc.

c. Exchange
- (*a*) prices — wholesale and retail.
- (*b*) sales — number of and amounts involved.
- (*c*) crises — financial and industrial.
- (*d*) failures — financial, commercial, and industrial.
- (*e*) etc.

d. Distribution
- (*a*) rents.
- (*b*) wages and methods of wage payments, real and nominal wages, etc.
- (*c*) profits — competitive and monopoly.
- (*d*) interest rates.
- (*e*) etc.

(3) Application to Governmental Discrimination and Policy.

a. The determination of the benevolent or malevolent effects of a given state policy.
b The determination of "fair values" and "reasonable returns" as bases for the exercise of administrative discrimination and the shaping of governmental policy.
c. The supervision of private business methods, looking toward the insuring of competition, the regulation of monopoly, the guaranteeing of favorable conditions of employment, etc.
d. The evaluation of properties as a basis for taxation, condemnation, and forced sale, etc.
e. The recording of domestic and foreign trade movements, estimating national wealth and its distribution, recording national progress so far as revealed statistically.
f. etc.

As a basis for the formulation of sound economic theory, the use of statistics and statistical methods is frequently necessary. Keynes has appraised this function admirably. The function of statistics is "first, to suggest empirical laws, which may or may not be capable of subsequent deductive

explanation; and secondly, to supplement deductive reasoning by checking its results, and submitting them to the test of experience."[1] Professor Moore's *Laws of Wages* is an excellent example of the use of statistics and statistical method in the development of economic theory. Stating his purpose, he says, "I have endeavored to use the newer statistical methods and the more recent economic theory to extract, from data relating to wages, either new truth or else truth in such new form as will admit of its being brought into fruitful relation with the generalizations of economic science."[2] This use of statistics and statistical method, while possessed of great possibilities in the hands of the well-trained statistical economist, offers few opportunities to the reader to whom this is addressed and shall not occupy a place in the discussion.

With this short introduction, the aim of which has been briefly to justify the submission of a volume on statistical methods, roughly to define the boundaries of the subject, and to suggest some of the broader topics to which statistical methods are applicable, we pass immediately, in Chapter II, to a consideration of sources and collection of statistical data.

References

King, W. I.—*Elements of Statistical Method,* Chs. II, III, pp. 20–39.
Bowley, A. L. — *Elements of Statistics,* Ch. I, pp. 3–13.
——— — *An Elementary Manual of Statistics,* Ch. I, pp. 1–6.
Pearson, Karl. — *The Grammar of Science,* Ch. I (Introduction).
West, Carl J. — "The Value to Economics of Formal Statistical Methods," in *Publications of the American Statistical Association,* September, 1915, pp. 618–628.

[1] Keynes, J. N., *Scope and Method of Political Economy* (2d Ed., revised), p. 338.
[2] Moore, H. L., *Laws of Wages*, p. 6.

CHAPTER II

SOURCES AND COLLECTION OF STATISTICAL DATA

I. Introduction

The first part of this chapter is devoted to a consideration of the chief governmental and private sources of statistical data which bear on business and economics. This is followed by a discussion of the tests to be applied to data in order to determine, among other things, whether they are biased, applicable to the case in point, exclusive or inclusive, whether the units are uniform, clearly defined and comparable, etc. The question of accuracy is next raised and attention given to statistical reporting, to the subject of errors and requisite accuracy, methods of estimation, etc.

The second part of the chapter has to do with the collection of data within and without business units. Attention is first directed to the preliminaries to the collection process, such as the availability of data, to the relation of those desired to others already collected, to the sanction back of the collection, and to the balance which must characterize the approach. The collection process is next described in detail. The discussion covers, among other things, the purpose and plan, sources, sampling, schedules and schedule making.

It is not our intention to chronicle in any complete way the great variety of types of statistical data that are now currently collected and published. Neither are we primarily

interested in cataloging the places where they may be found nor in passing judgment upon them. Such an undertaking would be as difficult as it would be tedious. We are interested, however, in citing certain typical data and calling attention to their elements of strength and weakness, but this is done almost solely for illustrative purposes and as bases for generalizations which it is desired to make. As it is no part of our task to compile a catalog of statistical sources, neither is it to our interest to make a compilation of statistical material which might be of use to the student or business man. A certain amount of the foraging instinct is presupposed on the part of the person who desires to use published statistics, and at least a general knowledge of what data are collectible on the part of those who are seeking original data.

It is entirely inadequate alone to know the sources of statistical data. Such knowledge is readily acquired. The ability to pass judgment on the worth of such data and to use them in a scientific manner is not easily gained. It is primarily the latter aspect of the problem in which we have interest. The former viewpoint in reality is subsequent to and conditioned upon the latter.

Statistics after all are in a large measure synthetic.[1] They are derivative in the sense that they express phenomena numerically, as they appear to an observer. Even the simplest facts enumerated require that conditions of identity be established. The counting of such a simple thing as ten

[1] "When we are investigating the nature and causes of things and events in the natural and social sciences, we are face to face with *facts*. In statistics about those events we are brought face to face with *syntheses*. The statistician must regard his figures as a sort of symbol, whose character and significance are more or less enigmatic; and he must diligently seek out all the probable causes of the facts he has symbolized before him, with a view to their scientific explanation." P. Coffey, *The Science of Logic*, Vol. II, p. 287.

bushels of wheat [1] would seem to offer no serious problem to any one advanced beyond infancy or the savage state. Yet it is not clear in this form what is meant by a "bushel," and, of course, wheat is not always a homogeneous commodity. Is the "wheat" dry or moist, spring or fall, hard or soft, etc.? Without now opening up the problem of units, and reserving for future treatment the human element in statistical studies, it is clear that a mere knowledge of sources does not make one a statistician.

II. Descriptive Sources of Secondary Statistical Data

By "secondary data" are meant those which have been collected, tabulated in simple or composite form, and made available for use, but which are removed one or more steps from the form in which they were reported and consequently do not show on their face the nature of the units employed, the purpose for which used, the treatments to which they have been subjected in analysis, etc. The term is used in contrast with "primary data," by which are meant those appearing in schedule or other original form, not having been combined into complex units, the characteristics of which may be understood by study. This expression suggests original studies; the former those which are secondary. It is secondary data which are generally used, — since they are readily at hand, — and unfortunately too often without a clear idea as to their merits for the purposes in mind.

The chief sources of secondary statistical data are the reports of public and private agents. These are either regular,

[1] See the interesting study by Boerner, E. G., "Improved Apparatus for Determining the Test Weight of Grain, with a Standard Method of Making the Test," *Bulletin* No. 472, United States Department of Agriculture, October, 1916.

irregular, or monographic in character. As examples of the regular type the publications of the United States Bureau of Labor Statistics, relating to index numbers of prices and to actual retail and wholesale prices, may be cited. Before 1907, this Bureau also published an index number of wages. Since that time, however, the wage data have been restricted to wage-rates in typical industries and have not been used to compute an index number for the country in general.[1] To this bureau we may confidently look for regular publications of current price and wage data. The thing which it is desired to emphasize now is the regularity with which the work is done and the continuity and substantialness which marks the policy under which the data are compiled. Other public organizations of similar character are the United States Census Bureau and the Department of Agriculture.[2] To each of these we are accustomed to turn for a great mass of statistical facts relating to conditions of production and ownership in manufacturing industries and to the development of agricultural resources, conditions of tenancy, etc. Bureaus of this type are constantly extending their spheres of activity so as to include in their publications the main facts of interest to the people as a whole and to certain groups in particular.

Within the states different public bureaus regularly issue statistics on a variety of topics. Some of these are of a high order of excellence and some are of questionable repute.

[1] In *Bulletin* 194, such an index is again computed, but is limited to union wages and the base is changed from 1890–1899 to 1907. See Rubinow, I. M., "The Present Trend of Real Wages," *Annals of the American Academy*, January, 1917, pp. 28–33.

[2] For an account of the United States Government's crop reports, see "Government Crops Reports: Their Value, Scope, and Preparation," *United States Department of Agriculture, Bureau of Crop Estimates, Circular* 17, Revised, pp. 8–26. This is reprinted in Copeland, M. T., *Business Statistics*, pp. 138–161.

There are also a number of regularly issued private statistical publications, not in the main duplicating, but rather extending, the work carried on by the public bureaus. Examples of these are the *Journal of the Royal Statistical Society*, which contains in the March number a résumé for the year of Sauerbeck's Index Number; *Bradstreet's*, *The Commercial and Financial Chronicle*, *The Financial Review*, *The Annalist*, all containing important price and market quotations. Current prices of commodities dealt in by boards of trade are published in the larger cities, and we are accustomed to turn to the reports of these organizations for detailed data.[1]

A tendency has recently developed for the Federal Government, particularly, to make extended statistical studies into special fields and to issue voluminous reports. Examples of these are the recent Immigration Reports, the Reports on Women and Children in Industry, the Report of the National Monetary Commission, etc. These, of course, belong to the public category. As examples of irregular private reports of a high order of excellence mention might be made of certain of the publications of the Russell Sage Foundation.

A third source of statistical data is the monographs which are constantly appearing on a great variety of subjects associated with economic and business topics. The mass of data collected in doctorate dissertations and in economic histories is often of a high order of excellence. Their chief function is to supplement the detail frequently omitted in

[1] For an account of the sources of statistics on produce markets, see Mudgett, Bruce D., "Current Sources of Information in Produce Markets," in *Annals of the American Academy of Political and Social Science*, Vol. XXXVIII, No. 2, pp. 104–125. This is reprinted in Copeland, *Business Statistics*, pp. 161–177. On some of the private organizations regularly collecting and issuing statistical data, see Parmelee, Julius H., "The Utilization of Statistics in Business," in *Quarterly Publications of the American Statistical Association*, June, 1917, pp. 565–576.

regular reports and to interpret those included. Excellent examples of this type are found in Smith's *The United States Federal Internal Tax History from 1861–1871*,[1] and Suffern's *Conciliation and Arbitration in the Coal Industry of America*.[2]

Besides these sources, mention should be made of the results of individual inquiries, which, while they are not necessarily carried on under competent supervision, nevertheless have considerable merit and may be used with discrimination by the student of economic topics. Other sources, containing material which may be characterized as hearsay and stray information, regularly or irregularly appear. Outside of the current financial sheet much of the statistical material appearing in newspapers must be looked upon with suspicion. As a source it is not to be relied upon alone. Its use must be prefaced by close scrutiny for accuracy of detail, completeness, and representativeness.

III. Tests to be Applied to Secondary Statistical Data

It is impossible to formulate a set of rules for the use of secondary statistical data which will serve as a complete guide under all circumstances. The best which can be done at this time is to point out some of the precautions which should be taken against too free use of this type of data and some of the consequences of ignoring them. The *first* consideration which should be mentioned is that of the bias or the unrepresentative character of the material. The old contention that "figures will not lie, but that liars will figure" is possessed of a substantial modicum of truth. When

[1] Smith, H. E., *The United States Federal Internal Tax History from 1861–1871*, Houghton Mifflin Co., Boston, 1914.

[2] Suffern, Arthur E., *Conciliation and Arbitration in the Coal Industry of America*, Houghton Mifflin Co., Boston, 1915.

prompted by motives to deceive, one has little difficulty in making out his case from data which if used otherwise would tell a different story. The bias may result by willfully eliminating part of the facts, by rigidly adhering to appropriate and clearly defined rules in the collection of material, but by basing comparisons upon insufficient data or by relating them to unrepresentative periods or conditions. If choice is made according to chance, an accurate picture or a trend may be shown from comparatively few data. If, however, choice is biased, an increase in the number of samples taken only tends to enlarge the amount of error. No use should be made of secondary data until the question of bias is settled. One should be fully cognizant of this point before analysis is begun.

A *second* consideration relates to the applicability of data to the problems being considered. Are the facts germane? Do the units of measurements in which they are expressed admit of use for the particular problem in mind? Many statistical data having only a general application may, if used with discrimination, substantiate or lend support to a contention which they would not be sufficient to uphold *de novo*. The bearing of these tests assumes importance only by detailed study of the uses to which one desires to put data and the conditions surrounding their collection. No single rule or principle is sufficient to cover all cases.

As to whether data are exclusive or inclusive is a *third* primary consideration. If it is desired to furnish a complete picture, then data must be scrutinized for their inclusiveness. If, however, the problem is merely to indicate a trend, then a different set of considerations maintains. If one were interested in the question of farm ownership and tenancy in a state, for instance, it would probably be necessary to study more than widely scattered sections since conditions are not

necessarily homogeneous as to the prevalence of ownership, nor uniform respecting the terms under which tenancy exists. Again, if the topics under consideration are types, amount, and economic status of immigrant labor, one would hardly be safe in restricting his study to a single port of entry. It might be possible by so doing to secure data which are typical of the total immigration, but more than typical facts are wanted. The problem suggests a quantitative and not alone a qualitative result. The same is true respecting studies of births, deaths, and accidents, etc. The recording of an occasional death by cause, an occasional birth, or a few of the serious industrial accidents is inadequate. What is necessary is the inclusion of all deaths by specific cause, the recording of all births, and a complete register of all accidents. Accident risks, for instance, cannot be properly determined unless all accidents occurring, the place where and the condition under which they happen, and the extent of disability, etc., are definitely known.

On the other hand, if all that is desired is to indicate the trend in a given set of facts it may suffice to take well-distributed samples. Undoubtedly, the phenomena of changes in prices can statistically be demonstrated without including statistics of all prices. If our problem is to measure changes in wholesale prices, this may be done by studying the prices of a comparatively few well-selected commodities over a period of time. The same may be said of prices of raw products or of goods in which the final consumer is particularly interested. The trend of the price of real estate, of stocks and bonds, may be indicated roughly by considering comparatively few but representative sales applying in each case. An illustration of this truth is found in the practice of real estate boards and tax bodies, in the use of sale statistics, to determine either the market or "true

value" of real estate. The chief consideration is the representative character of the samples. Wage increases or decreases may be shown by a process of sampling, providing the samples are chosen with discrimination. If it is desired, for instance, as evidence of the value of a piece of property, to enumerate the number of people who pass it, it is sufficient to include relatively short periods typical of both rush and slack hours for representative days. The enumeration of the entire number of people of all classes for an extended period is unnecessary. Likewise, the scale of rents in a given district may be determined with sufficient accuracy for commercial purposes by considering rents of representative houses. It is not necessary to include all houses. Care must always be exercised, however, to see that the sampling, howsoever carefully made for purposes of original compilation, is suitable for the purposes in mind. It may be formulated as a general rule, that the more nearly all data are included the less is the likelihood of bias controlling, and the more readily can they be converted to a particular use. Under such circumstances the particular facts desired may more easily be chosen and extraneous ones eliminated. Again, however, nothing better than general principles can be laid down as a guide in the appropriate use of secondary material. Discrimination, caution, and eternal vigilance are essential prerequisites to scientific study and to the formulation of valid conclusions.

As to whether units of measurements are *simple* or *composite* is a *fourth* consideration. By simple units are meant those in which one determining consideration is prescribed. Most statistics of enumeration employ simple units, as for instance, where persons, animals, acres, buildings, passengers, stocks, deaths, laws, sales, etc., are merely counted. In statistics of this type the disturbing elements due to inac-

curacies in the units are reduced to a minimum. Nothing of course is said concerning the accuracy with which units are defined, the rigidity with which definitions are followed, nor the accuracy with which enumerations are made, but only of the fact that the presence of a *single* disturbing cause associated with units normally guarantees against the presence of greater or as great a degree of error than would be associated with conditions when units, and hence statistics, are composite in character. Such a unit as a "farm" might easily be defined and the statistics of farms readily be understood. When, however, the limiting expression "improved" is added to this unit, the scope of the definition and its application have been materially restricted, and an additional element introduced into which error may enter with the same readiness as into the other portion of the combined unit. Likewise, in statistics of "daily wages," of a "fair return," there is introduced possibility of error from definition, not only from one but from two sides. Crops in bushels or in acreage may readily be determined; the "normality" of these crops, however, raises other problems and calls for superior statistical organization and for a much greater exercise of judgment. As these additional considerations enter, occasions for error and bias crowd in, and it is these conditions to which attention is drawn in distinguishing between simple and composite data.

Numerical data may be expressed in the form of ratios, or relative numbers. These are known collectively as *coefficients* and imply definite relations between numerators and denominators. A coefficient should be assignable to the conditions which make it possible, or in the words of Bertillon, "always compare effects to the causes producing them." One would not relate the number of deaths from spinal meningitis to the whole population, nor compare in

this respect populations of entirely different age composition. Neither would one compare the number of industrial accidents for similar plants where the hazard or exposure in terms of man-hours and machine-hours is widely different. Likewise, statistics of the number of farm accidents should not be related to the total number of farm employees, but only to the number employed in occupations producing the accidents. The number of accidents occurring in the mining industry would seem to stamp it as highly dangerous, yet this is noticeably true only when the accidents are related to the types of occupations in which the hazard is exceptional.[1]

Loose thinking always results when effects are not related to the specific causes producing them. Long hours, poor ventilation and light in factory or mill are often assigned as the causes of occupational disease and laws are passed to correct the evils; yet it is not always clear how much of the result ought not to be assigned to conditions of home life, intemperance, etc., things only remotely associated with or entirely disassociated from the occupations *per se*. In each case responsibility can be assigned only after investigation and after each effect is related to its specific cause.

It is not a sufficient justification for the violation of this principle to maintain that in economic life effects are rarely if ever to be attributed to single causes, and therefore all effort to allocate the responsibility is useless. The statement is true but the inference does not follow, and its truth only calls attention to the extra care necessary in the use of economic and social statistics before conclusions are drawn from them and policies mapped out upon them. Here again, the best that may be done is to call attention to this

[1] For a more complete discussion of *Units of Measurements*, see Chapter III, *infra*.

important fact and leave the investigator, thus warned, to make application of it in each problem considered.

A *fifth* consideration is that the use of data is conditioned, among other things, upon the *accuracy with which reported, the accuracy with which determined*, and *the accuracy of determination*. Each of these topics requires brief consideration.[1]

The accuracy with which data are reported and collected depends in large part upon the character of the informant, the nature of the records kept, the type of questions asked, and the care used in answering them. If difficult and unfamiliar questions, or questions which in any way incite distrust or suspicion, are asked, the answers are likely to be either incomplete, brief, non-committal, full of error, or purposely evasive. The problem largely turns on the question of reporting. Age, for instance, may be accurately known, but falsely reported. Wages may be known and yet not reported simply because of suspicion of the use to which the data will be put. Moreover, even in cases where there is no reason for falsely reporting, liability of error in tabulation is always a factor to be considered. The amount of accuracy carried into the final returns depends upon the care used in editing, and the general manner in which the tabulations have been made. Devices permitting clerical accuracy have been pretty well perfected and are now in common use. Glaring errors may be detected by an analysis of the data themselves. It is seldom necessary, however, to check the numerical computations of reputable statistical publications; it is always necessary to satisfy oneself of the character of the primary material which is the basis for secondary tables.

[1] For discussion of similar points respecting wage data, see Chapter IV, "Types of Secondary Wage Data."

On the other hand, data may correctly be reported but the report itself be inaccurate because the answer is wrongly determined. Much of the data, until recently, respecting causes of death fall under this head. No necessary difficulty is experienced in reporting, but only in determining the precise cause, or in calling by the same name the same thing. The necessary corrective is, of course, a standard classification of *causes of death*, and this we now possess for the so-called registration area of the United States. Likewise, statistics of occupations in the United States suffer greatly from the lack of a standardized nomenclature. Identical occupations are called by different names; things which are equal to the same thing, in reality, are not equal to each other in name. As a basis for the determination of occupational risk, for the development of schemes of accident compensation or insurance, they are almost worthless. Fortunately, we are now making some progress toward uniformity of occupational naming. Here, as in the former consideration, the personal equation is important, but more often the real source of trouble lies, as in the instances cited, in the nature of the problem itself.

Statistics of capital employed in manufacturing industries, as reported by the United States Census Bureau, suffer much because of the inaccuracy with which determined. The definition of capital for statistical purposes offers the first difficulty. Even for detailed analysis authorities are not agreed as to what should be included as "capital." The reasons for including or excluding different categories vary and are of different force in different industries, or in the same industry under different conditions of management and forms of business organization. For census purposes such a unit must of necessity be used with little more than a semblance of exactitude, and, of course, the statistics col-

lected are very little better than rough guesses. The same considerations apply to "value of products," "cost of materials," "expenses," etc. The difficulty is not necessarily one of error in reporting (yet undoubtedly this is an important factor) nor in the accuracy with which such facts *might be* determined, but rather with the accuracy with which they *are* determined under the conditions of collection. If nothing more is desired than an indication of trend this may be secured in cases where complete accuracy of detail is wanting, providing errors are distributed uniformly about the average and tend to correct each other, and where sampling is representative. These conditions, however, so seldom maintain (never in the last instances cited) that data compiled ostensibly under these considerations must be used with great care and circumspection for any use where accuracy is important or where vital issues are involved. It is painful to see nice distinctions and weighty conclusions rest upon such questionable support!

On the other hand, secondary statistical data are frequently compiled where absolute accuracy of determination is impossible and where no pretense is made toward completeness. The data at best are estimates. At present no statistical machinery and data are available for an accurate determination of the amount of gold-producing ore in the United States; of the amount in horse-power of our water power resources, or of the amount of standing timber existing in the United States.[1] Absolute accuracy is not necessary and no pretense is made of its realization. Of course, there may be accurate and there may be inaccurate estimates,

[1] See the interesting report on "The Lumber Industry, Part I, Standing Timber," by *The United States Bureau of Corporations*, 1913, where methods of estimating the amount of standing timber in various districts and for various woods are described and criticized, pp. 7–10, 45 ff.

and it is always incumbent upon him who uses them to choose those which, all things considered, seem best to meet the requirements which they should possess and to use them as *estimates*. Essentially accurate conclusions, of course, may be drawn from rough estimates, but in their use the element of danger is so great that caution should always accompany their employment, and sound judgment constantly be invoked to guard against false conclusions being drawn from them.

Moreover, not all phenomena allow of statistical measurement. Numerical frequency may be of no real nor vital significance. The devotion of a people to a principle of right or justice can hardly be measured by the number of those who find no occasion to violate it. Regard for law and order may not be measured by the number of people who remain out of jail. Conversely, the disregard for law is not fully measured by the number of arrests and convictions for a given period. The degree of insanity is not necessarily measured by the number of commitments to insane asylums together with the number of occupants of such institutions. The sacredness with which the marriage institution is regarded is not accurately reflected by the number of divorces granted, nor respect for higher education alone by the number of students enrolled in institutions of collegiate and university rank. It is an error to expect statistical data alone to answer these questions, and it is even a worse error solely to base conclusions respecting them on data which are now extant.

Not less important than the element of accuracy is a *sixth* consideration, viz., the homogeneity of conditions which the data describe. If violent changes of methods of doing business have resulted during a period of time, and the corporate form of organization has become more common

because of the relative size of the business unit, then it would be inaccurate to base conclusions respecting the proportion of business done under this type of organization at two periods where all business is taken as the basis of comparison. If "future" transactions, in a given market, are supplanting "spot" transactions, and the substitution has caused prices to rule higher or lower, then prices to-day may not be compared in this respect with those characterizing a period when such methods of dealing were not indulged in. If prices to-day are influenced by the practice of retail dealers "protecting" manufacturers by refusing to give price concessions, then present prices are not fully comparable with times when such conditions did not maintain. If price levels are to be compared, it is unfair to make the basis of comparison prices of commodities bought in small quantities with those paid for in wholesale lots. The conditions are not equivalent, and comparisons are invalid until they are reduced to a common denominator. Prices expressed in a depreciated standard can be compared with those made on a gold basis only after a conversion of one has been made into terms of the other.

Not only may statistical data be descriptive of non-homogeneous conditions (and this fact not be revealed), but they may also greatly differ in composition at different times. Reporting, editing, tabulating, and analyzing may be of widely different degrees of excellence. New forces may have been given recognition, different emphasis may have been placed on different things, different definitions may have been insisted on, new units of measurements or modifications of old ones may have been employed, wider or narrower fields may have been covered, the proportional elements used to make up a total may have changed materially, etc. The presence of these and similar conditions makes

comparisons over long periods difficult, if not exceedingly dangerous. The desire for "comparability" often becomes the controlling factor in statistical computation, and serious omissions, strained interpretations, etc. (all important in the use of the data for a given time) countenanced in order to preserve it. The retention of the "capital" inquiry, in all its crudity, in the statistics of manufacture by the United States Census Bureau is largely out of consideration for the "value of comparisons." The omission until recently on the part of the United States Bureau of Labor Statistics of fifteen commodities formerly used in the computation of an index number of retail prices, raises at least the question of the possibility of comparing the figures before 1907 with those since that date.[1] The various definitions of a "farm," or of an "establishment," or of "manufacturing" used by the United States Census Bureau at different times, make hazardous comparisons over an extended period. Exports and imports, whether expressed in quantity or in value, must always be interpreted in terms of the units of measurement employed.[2] The student should always go behind

[1] The lack of comparability has been definitely asserted by the Commissioner of the Bureau of Labor Statistics. "Some Features of the Statistical Work of the Bureau of Labor Statistics," Royal Meeker, Commissioner, *Publications of the American Statistical Association*, March, 1915, pp. 431–441.

[2] Most interesting discussions of the difficulties of making international comparisons of import and export statistics, and of the imperfections of our own import and export statistics, are contained in an article by Frank R. Rutter on "Statistics of Imports and Exports," in *The Publications of the American Statistical Association*, March, 1916, pp. 16–35. Apropos the topic here under consideration the following extracts are of interest:

By virtue of a law passed in 1893 the agent of a railroad company carrying goods to a foreign country by land was made punishable to the amount of $50 for failure to present a manifest to the collector of customs. "The effect of the change in law is reflected in the exports through Buffalo to Canada. From less than $500,000 in 1890 the figures jumped to over $4,000,000 in 1895." *Ibid.*, p. 20.

On the matter of units of measurements and classification, the following

the printed figures and make sure of the units, their interpretation, and the weight assigned to the different factors

quotation is of interest: "The greatest need for the expansion of the classification is found in the case of exports. The most detailed classification of exports now covers less than 600 items, while in the imports for consumption there are about 3,000 distinct items. The chief preventive of an increase in the number of items is the indefinite character of export declarations. So many articles are described merely by general terms that it is out of the question to separate articles frequently of much commercial importance.

"Defects in the present classification, aside from its incompleteness, are the incomparability of the import and export schedules and the failure to conform to current commercial terms. The latter defect is due to the preservation in the tariff of many terms now obsolete, and the necessity of having the statistical classes follow closely the tariff items." *Ibid.*, p. 26.

On the definition of "imports" the author says:

"What is generally understood by the term 'imports'? Legally, an article is imported when landed, whether for immediate consumption or for storage in bonded warehouses. From an economic point of view, however, bonded warehouses may well be regarded as foreign territory. The door of the bonded warehouse is really the economic frontier of the country.

"Since the United States is not a large reëxporting country, the difference between 'imports' and 'imports for consumption' is largely one of time. The instances in which goods are exported from warehouses are few as compared with the instances in which after the lapse of time goods are entered for consumption within the country.

"Perhaps the distinction is most clearly brought out by an illustration. While the last tariff was under discussion wool in large quantities was landed at our ports and stored in bonded warehouses until December 1, 1914, when it could be withdrawn without payment of duty. Was such wool really imported when it was landed or when it was removed from the warehouse?

"On the export side we have a clear distinction between domestic exports and foreign exports. On the import side imports for consumption are most nearly comparable with domestic exports, yet not fully comparable, since free goods are not generally warehoused and may be entered for consumption although intended for reëxportation. To be strictly accurate, dutiable imports for consumption should be compared with domestic exports and free imports with domestic and foreign exports combined." *Ibid.*, p. 28.

"Perhaps the most striking instance of the unfortunate result of our method of valuation is seen in the import prices of rubber. Notwithstanding the improvement of plantation rubber, Para rubber is still quoted at a slightly higher price. In Brazil, however, there is a heavy export duty, which constitutes an important element in the price. This duty is not included in our statistical valuation with the result that the value of India

in the composite group before he hazards detailed comparison or arrives at conclusions.[1]

IV. Considerations of Importance Prior to the Collection of Data

Before undertaking a statistical study it is essential that the problem be studied in order to determine the possibility of the statistical as contrasted to other approaches. All problems do not lend themselves equally well to numerical treatment. Indeed, many questions are so affected by ethical, moral, and religious considerations that they do not admit of statistical interpretation.

If it is decided that the problem possesses statistical merit, among the important things to be considered before actual collection of data is undertaken is the availability of the facts desired. Not infrequently data relating to a given phenomenon *exist* but are not *available*. This condition may result from the fact that records are imperfectly kept, that data are so meager and so widely distributed, or scattered over so long a period of time that the expense involved makes collection impracticable. In the case of industrial occupations frequently we have only the trade name, or the trade processes, available and it is difficult to reduce to a uniform nomenclature the reported facts as a basis for any valid conclusions. If data desired are available, they still may not be in a form which will permit of their being directly applied to the problem at hand. Conversion of the units may be necessary. This frequently requires technical knowledge

rubber imported from Brazil during the fiscal year 1914 averaged only 40 cents a pound, while the import value of that from Ceylon averaged 60 cents a pound." *Ibid.*, p. 30.

[1] Bowley, A. L., "The Improvement of Official Statistics" in the *Journal of the Royal Statistical Society*, September, 1908, Vol. 71, pp. 461–469 particularly.

and in many instances the use of unwarranted discretionary power.

Besides availability, the relationship of data to be collected to complementary and supplementary facts already collected or possible of collection should be considered. This suggestion has to do with the necessity of correlating existing statistical material rather than with the technique of actual collection. Yet it is intimately connected with the latter. Indeed, the type of data already available may be the dominating factor in determining the new line of statistical approach. To duplicate work already done is justifiable only when it is felt that existing data are incomplete, unrepresentative, or in some other respects inadequate or unsuited for the uses to which one desires to put them. The aim should always be to supplement, to carry one step further, to make function the data already possessed. Too frequently statistical studies both of students and of statistical bureaus are uncorrelated. They stand out as independent efforts, throwing little light upon problems to which they are addressed largely because they do not form a necessary part of a single and comprehensive program. They begin and end as independent, uncorrelated efforts.

An illustration respecting public bureaus will serve to bring out the importance of this consideration. The statistical bureaus of some of our leading states collect from one to three, or possibly four, important types of data upon the subject of unemployment. Taking Massachusetts as an example, we note four types. The first is the one on unemployment due to lack of work, lack of material, strikes, lockouts, etc., regularly collected from trade unions. These data apply only to *union* conditions. A second type, rather upon the subject of employment than unemployment, is the data on the average number of em-

ployees by months reported by manufacturing institutions to the Department of Manufacturers in the Bureau of Statistics. A third type, more local in its character, is that regularly collected by the Public Free Employment offices. The facts in this case relate to the applications for employment and, of course, cover both union and non-union employees. A fourth type exists in the form of data regularly collected concerning accidents and compensations for accidents by the Industrial Board. These types of information, although separate and distinct in character, undoubtedly throw considerable light on the subject of unemployment, and if correlated would bear even more strongly upon the problem. Up to the present time, however, the collection of each type of information has been considered chiefly as an end in itself, and no systematic attempt has been made to correlate the material collected.

The lack of coöperation and the overlapping of function and output of American statistical bureaus generally are appalling. Respecting the national government it has been suggested recently that there be created "The Office of National Statistics" to act as the coördinating unit among the "twenty-nine branches of government" now issuing statistics with the "inevitable plenty of wastefulness and duplication." The lack of coöperation between bureaus of the Federal Government may be shown by the following illustration:

The law providing for the taking of the United States Census makes it obligatory upon every manufacturer to supply Census data, and stipulates that the information furnished "shall be used only for the statistical purposes for which it is supplied. No publication shall be made by the Census Office whereby the data furnished by any particular establishment can be identified, nor shall the Director of

the Census permit any one other than the sworn employees of the Census Office to examine the individual reports." Precautionary measures are undoubtedly necessary to guard against publicity of individual returns to the detriment of those involved in competitive industry, but it does not seem necessary and reasonable for this bureau to make a fetish of this restriction and in a measure to defeat the purposes of other departments of the government. This publicity provision is now (1915) so narrowly interpreted as to preclude other departments of the government from even securing a list of the names and addresses of the manufacturers to whom the Census sends schedules. Let us see just how narrow such a policy is, and some of its consequences.

On any given Census year the chief sources of materials for names, addresses, nature of business, etc., are the schedules on file for the preceding census. These must be corrected and supplemented from trade directories, telephone books, gazetteers, etc. To correct a list for the United States is an enormous task, and if done by one department of government its duplication by others would seem unnecessary. The Census Office, however, so narrowly interprets the confidential features of the law as to refuse to furnish the list to the Trade Commission, notwithstanding the fact that only by the merest chance could the Commission, if it desired, clearly distinguish the names and addresses for those cases in which these facts were not generally known, whether supplied from old schedules or from directories. The Census has the necessary facts and organization for compiling a complete list at a low cost, yet after it has compiled the data for administrative purposes, their use by other departments within the national government is refused.

The result is as follows: Within the Federal Government (not to speak of the state departments to which such a list

might be furnished upon some reasonable basis) there are, among others, the Census Bureau, the Bureau of Labor Statistics, the Trade Commission, the Children's Bureau, all requiring as a first condition to the administration of law a list of manufacturers, traders, mercantile concerns, etc. The *reason* for desiring the list *varies* with the departments; the *necessity* for the list is *common*. The Census Office in the face of this common need refuses the information under the flimsy pretext that the matter is "confidential." Such lack of coöperation, whether resulting from the provisions of law or from the short-sighted policy of the administration, should not be allowed to endure.

Only recently have there been any serious attempts to correlate and standardize the statistical work of the several states. The passage of workmen's compensation laws by a majority of the industrial states has demonstrated the necessity of the adoption of a definition of an industrial accident, the use of uniform report blanks, and of uniform methods of tabulation of accidents. Under the leadership of the United States Commissioner of Labor Statistics, and in coöperation with the statisticians of the bureaus of the states affected and the liability insurance companies, there are gradually being developed uniform standards in definitions and treatment of industrial accident statistics.[1] Until these are in use it is impossible to reduce industrial statistics to a comparable basis and to calculate the degrees of hazard accompanying occupations either for purposes of workmen's compensation or state insurance.

There are instances, likewise, in which the states are coöperating with the Federal Government in the compilation

[1] The progress in this line is conveniently summarized in "Industrial Accident Statistics," *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 157, March, 1915, Washington, D. C., 1915.

of statistical information. Massachusetts, respecting her manufacturing census; Ohio, respecting union rates of wages; Wisconsin, respecting labor in canneries; Illinois, respecting industrial disease; Indiana, respecting female wage earners in mercantile establishments, — are cases in point. There are similar instances in which the work of the states, if not completely duplicating that of the Federal Government, is done in ignorance of it. New Jersey's cost of living studies and the refusal of numbers of states to accept the provisions, respecting reporting of births and deaths, established by the Department of Vital Statistics, Bureau of the Census, are conspicuous.

No attempt is made to compile a catalog of the multitudinous points of statistical contact between the statistical departments within the national government, within the states or other divisions, or between the departments and the several jurisdictions. Neither is it the intention to enumerate the instances in which coöperation is effected or in which it is ignored. Conspicuous instances of coöperation and its absence stand out, and these have been mentioned for the purpose of calling attention to a problem the study of which in the United States has been sadly neglected.

The examples cited will suffice to bring to the attention of persons and bureaus intending to make statistical inquiries the necessity of studying the field so as to become acquainted with what has been, and is being, done in order more properly to make the facts collected supplement rather than duplicate matter already collected. By this simple expedient many inquiries, which in themselves will be fruitless because of lack of time and money, may be avoided, and real contributions made by gathering additional evidence on single or closely related phases of topics or by correlating material already at hand. One cannot legitimately object

to the industry displayed by modern statistical bureaus, in collecting facts; but severe criticism of the disposition to consider collection as an end and to leave untouched any contrasted and correlated use of the material is frequently justified.

Another consideration of importance prior to the actual collection process should be mentioned. Most public agents are possessed of mandatory power. They may compel answers to be made to the inquiries submitted. This power normally does not extend to private individuals and its absence in most instances is a real handicap to effective inquiry. It is, however, sometimes possible for investigators, through contact with informants, and by cultivating their good-will, to develop in them a feeling of obligation to report, which more than compensates for any lack of mandatory power. So far as public statistical organizations are concerned, conspicuous instances where a feeling of obligation to supply information has been well developed are the cases of price reporting to the United States Bureau of Labor Statistics, and the reporting by unions of the conditions of employment to the Bureaus of Labor Statistics in Massachusetts and in New York.

By cultivating the good-will of informants, these bureaus have been able to enlist their support, so that at the present time they receive excellent reports with little actual inconvenience and cost. Various ways are open for securing their interests and good-will. One approach is through a guaranty against an abuse of confidence. Sometimes it is accomplished through assurances that statistics desired apply to the group as a whole, and when compiled will be supplied gratuitously to all those who have contributed to their collection. Sometimes appeal is made openly to the feelings of state or local pride, as, for instance, in the collection of

statistics of manufactures in New Jersey. In this instance the bureau inserts a provision in the manufacturer's schedule to the effect that in case answers are not made the returns for the state will be deficient, and that relatively New Jersey's showing will be less favorable than that made by other states. Another way of gaining the confidence of the informants is by studying their interests and by cultivating their good-will by correspondence. This method is being used effectively in Massachusetts, where bureau officials are careful to indicate by semi-personal letter the value to the informant and to the public generally of data to be collected, and the importance of answering specifically and promptly the inquiries made. Where mandatory power exists it is not an uncommon practice for statistical bureaus requesting information to quote the terms of the law, and to indicate the penalties attached to failure to live up to its provisions. This method, however, should be used with discrimination inasmuch as it may tend to incite a spirit of distrust and opposition rather than of coöperation.

Private individuals, as contrasted with regularly constituted authorities, may always be said to be in a disadvantageous position in this respect in the collection of data. The limitations under which they operate should be clearly kept in mind in order to guard against a too sanguine belief in the efficacy of individual effort. Too great confidence as to the outcome of a given undertaking generally characterizes the efforts of the inexperienced.

Still another consideration is of importance preparatory to the collection process. It is necessary to know the types of informants to whom appeal must be made. If they are ignorant, indisposed to appreciate the significance of the problem under study, or to oppose its continuance, if they are inclined to look upon everything as inconsequential and useless,

little weight can be attached to the answers given. The considerations named above in cultivating a personal acquaintance apply here. An investigator, however, should be cognizant of this limitation, and should consider it as a preliminary fact to be given attention before an extensive statistical study at first hand is undertaken. Likewise, the time, the money, and organization available should be considered. Data may exist, informants be ever so willing to supply them, and yet the actual consummation of a task be impossible because of lack of funds, time, or organization. Few people not accustomed to planning statistical work clearly realize the time, energy, and expense involved in a thorough statistical investigation.

V. The Collection of Primary Statistical Data

1. *Purpose and Plan*

In the actual collection process the first and foremost considerations are the purpose and plan. These should be outlined clearly both as to direct and indirect implications. The scope of the problem should be thoroughly understood and the primary and secondary considerations bearing upon it clearly realized. The limitations of the statistical approach should constantly be held in mind. All units to be employed in actual measurement should accurately and unmistakably be defined and the problem, so far as is possible, viewed from beginning to end. Only by so doing is it possible to provide in advance for all contingencies that may arise and to make an adequate statement of the case. The ability to do this comes only with practice, but the necessity of its being done is no less real by virtue of this fact. The problem of adequately and fully stating the purposes of statistical studies is held to be so important that

much of Chapters III and IV is devoted to a discussion of it for typical cases.

2. *Methods of Collecting Data (Descriptive)*

After having clearly outlined the problem and developed the plan to be followed, three methods of collecting data are available. The one or ones used will depend, of course, upon their appropriateness to the purposes in mind. The present treatment of these types is purely descriptive and does not attempt to outline all of their peculiarities and adaptations. *First*, recourse may be had to official records. In the case of business houses, undertaking statistical studies from data in their own records, the process of collection might, perhaps, more properly be called "assembling material from records." In many cases, no doubt, considerable adjustment in the types of records, and in the manner in which facts are reported, is necessary before they can be made available for summarization and analysis, but in these cases the presumption is that after the preliminary work is done — and oftentimes this is a real and vital part of the problem — that the remainder, so far as the collection of material is concerned, is largely a question of transcribing data. Motives for withholding part of the facts, of misstating those given, or of blocking the study with the purpose of defeating it, are not presumed to be present, since the purpose of undertaking it is to throw light on the relative efficiency of methods pursued and to point the direction for possible changes in policy, organization, etc.

Moreover, the conditions for the operation of personal bias, desire to make a case, reliance on incomplete returns are reduced to a minimum. The position is not taken that data available in returns currently collected or in those

which may be secured are always adequate, particularly when the purpose is indefinite, as it almost always is, in case it is not undertaken by some one especially trained for such work, but that those collected under these circumstances do not present the difficult problems which confront the statistician who comes to the work from the outside with no sanction except that of an impersonal government, a loose organization, or his own good intentions, and without the tact to enlist the sympathy and coöperation of those upon whom he must depend for success.

It is, of course, true that most smaller business houses do not understand the uses to which data in their own records can be put, and consequently do not have satisfactory statistical records. Moreover, those who appreciate their possible significance have considerable reservation about giving over to a separate department the function of informing others of the weak places in their organizations and of the losses which could be prevented and savings made. "Statistics" are in ill-repute and largely so because they are considered either in themselves infallible or fallible, — depending on whether they show the right or wrong thing, — or are used in an unscientific manner and, as a consequence, are not reliable. There is almost as much science in the way statistics are collected as there is in the subsequent interpretation of them, but this truth is almost the last to be recognized by the inexperienced.

If the agent securing data is outside the organization, records may be furnished in the original or their contents transcribed. If transcribed, this may be done either by the informants or by the agent. The former method is expeditious but liable to abuse. In some instances requests may be ignored or answers purposely misstated in order to deceive. Without an adequate check upon the information

furnished this method cannot be advocated as wise for general adoption. Examples where informants supply material from formal records, and still a reasonable degree of accuracy is obtained, are the reports of accidents to various state compensation bureaus, and the reports of manufacturing statistics made to the Division of Manufacturers in the Bureau of Statistics, State of Massachusetts. On the other hand, instances are common where informants supply material which is grossly inaccurate. Accident reporting in New Jersey may be cited as an example. In the former case (Wisconsin may be used as an illustration), inasmuch as both insurance companies and employers are required to report upon accidents to the Industrial Commission and all employers are required to be insured, essential completeness and accuracy of accident reporting are guaranteed. In Massachusetts, manufacturing statistics have been collected from representative concerns for a great number of years and records are available for intimate comparison from one period to another. Under such conditions it is almost impossible that material error shall characterize the figures, particularly in view of the care exercised by the bureaus in their compilation.

Where schedules are used and informants are required to fill them out, the necessity for detailed descriptions of units is often so great in spite of extreme cautions that serious errors creep in. Long explanations cannot conveniently be made upon schedules, and it is impracticable to accompany them with elaborate instructions. Only in cases where obligation is felt on the part of informants to answer questions, or where answers may adequately be checked or given under supervision, as, for instance, in the statements of expenditures in a recent study of working women's budgets in Ohio, can complete reliance be placed in information sup-

plied by schedules which informants themselves have filled out. In the investigation into "Wages and Regularity of Employment in the Cloak, Suit, and Skirt Industry," etc., in New York, information supplied upon 1429 schedules filled out by the workers and gathered by the shop chairmen, was found to be "so full of errors that they were discarded as entirely unreliable." [1]

An alternative method is for an agent or representative to transcribe the records. This is expensive but conducive to uniformity and accuracy. It is, however, not carried on to any great extent in large-scale investigations. A conspicuous instance where it is followed is in the statistics of cities, published by United States Census Bureau. By the use of agents, the Census Bureau is able to convert dissimilar accounting systems to an essentially uniform basis and to publish in most respects comparable statistics. This method has been followed to some degree in the collection of statistics of manufacture by the United States Government. In special investigations, such as those made by the Bureau of Corporations into the Petroleum Industry, the Tobacco Industry, the International Harvester Co., *et al.*, it is the rule.

A *second* general method of securing data may be described as the *process of counting*. Obviously, enumeration in some form is involved in all methods of collection. It is fundamental to the study of statistics. But in this connection enumeration or counting is used in a narrower sense with the idea of suggesting the process of initial count or tally. Where it is used, records do not generally exist to which direct appeal can be made; or if they exist, they are not currently corrected and it is desired to get more recent figures. The distinctive character of this process will more

[1] *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 147, p. 14, Washington, D. C., June, 1914.

readily be appreciated if examples in which it is used are cited and comments made upon them.

Probably the best example of a statistical study in which the process of counting or enumeration is primary and where it is most severely tested, and its limitations most emphatically revealed, is the United States decennial population census. Similar but less conspicuous examples are the regular or irregular state or city censuses of population. The surplus of births over deaths, together with the surplus of immigration over emigration, are the sources making for an increase of our population. Reasonably accurate statistics of births and deaths are restricted in the United States to the so-called registration area. Statistics of immigration and emigration are reasonably accurate for the country as a whole. Statistics of distribution of immigrants more accurate than possibly the state to which they *declare* they are bound, or of the origin of the emigrant, more definite than his last place of residence, we do not possess. Little or no record is kept of migratory movements of population within the country. The result is that for statistics of population we must chiefly rely on the decennial census made by the United States Bureau of the Census, and for the interdecennial years upon the state censuses or the estimates made by reputable statistical organizations.

The actual enumeration of the population of 100,000,000 people in a district as large as the United States is a gigantic undertaking. Divorced from the tendencies for districts to exaggerate their figures and for the enumerators to pad their lists in order to increase their remuneration, the difficulties are almost insuperable. Coupled with these conditions, and serving the political purpose which a census does, as an actual enumeration or count, little value so far as absolute or even near accuracy is concerned can be attached to it.

With the reasons for this state of affairs, attributable as it is to the method of appointment of enumerators, to the inherent bigness of the task, to the divided duties of the enumerators between a population census proper and an agricultural and occupational survey, to the political purpose which it serves, etc., we are not here particularly concerned. Our chief interest is in stating the method employed in the enumeration rather than in analyzing the trustworthiness of the data collected. Questions involving the determination of legal residence, the treatment of floating population, of people in transit from place to place, etc., are involved in the process of counting. Questions relating to classification, depending as the latter does upon race, conjugal condition, nativity, etc., are not peculiar to the problems of enumeration, but are present in all processes of collection. They involve the formulation of accurate definitions of the units employed, and rigid adherence to the conditions laid down.

In the case of the population census, partial checks on the accuracy of the count are found in the preceding censuses, in the records of deaths, births, immigration, emigration, and in the fact that normally the distribution of age and sex classes is essentially uniform from period to period (this relationship is somewhat disturbed in the United States by the influx and egress of mature male immigrants). These checks, however, valuable as they are to keep in bounds of reasonable inaccuracy the results of the canvass, do not lessen the inherent difficulties even under the best of conditions, of counting large aggregates even with approximate accuracy. The frequency of contested elections in cases where crookedness is admittedly absent, furnishes another evidence of the inherent difficulties in correctly counting large aggregates. As a generalization, however, it may be maintained that the difficulties experienced in enumeration are not so much to be

attributed to the inability of the mind to comprehend large aggregates as to the use of inadequate statistical organizations and the not infrequent desire to actually misstate a fact or misinterpret a condition of affairs.

The *third* method or process of collecting data is that of *estimates*. These may be made on the basis of formal records or of enumerations without records. They may be made on the basis of direct material, as when expectancy of death (life tables) is based upon the number and conditions of deaths. They may also be made from allied material, as when call-loan rates of interest are estimated on the basis of bank reserves, the net interior movement of money upon the size of crops, the trend of business on the combined factors making for business distrust or confidence, or the probable price of corn from the price of wheat. Indeed, in the business world most dealings are hazarded upon the ability to foretell the most probable results from a given set of conditions. Market prices of cereals are, in large part, a reflex of the likely condition of croppage during the subsequent six or twelve months balanced over against the likely conditions of demand; prices of securities are based upon an estimated earning capacity of the properties floating them; increases of investment are hazarded upon a continuance of favorable trade conditions or the favorable disposition of the legislature.

Much of the statistical data regularly compiled on the agricultural outlook, on the depletion or conservation of resources, upon national wealth and its distribution, upon the benevolence or malevolence of a given state policy toward business and industry, or the likely consequences of the adoption of a régime of Socialism or government ownership, upon the deleterious effects of a given work policy or condition upon the laborer, have nothing more solid at base than crude

estimates. Some of the data are sufficiently accurate for all practical purposes, are compiled under conditions which tend to give them real value, since absolute accuracy is unnecessary, and may serve as bases upon which to formulate a policy or launch a program. Such, undoubtedly, is true respecting the data issued by the Agricultural Department at Washington on the condition of crops, on the acreage of cereals, etc. Absolute accuracy is not required, and the amount of error, tending as it does widely to distribute itself and to remain essentially the same from period to period, is not a seriously disturbing factor. The same in part may be said concerning receipts and expenditures, earnings, tonnage, etc., of business units so currently and confidently estimated in business life.

On the other hand, estimates made respecting conditions which constantly change, and upon which adequate data as guides do not exist, or which in themselves are impossible of determination, have serious limitations. Too free use should not be made of them in shaping governmental or business policies and in questioning social and economic institutions. The estimated amount of arable land in the United States is materially increased by the completion of irrigation projects and the perfection of dry farming methods. Power sites are materially increased in number and value by the perfection of high power transmission lines, and the available supply of precious metals by the discovery and use of the cyanide process for separation of gold from crude ores. The estimated coal supply takes on new significance in the light of recent discoveries respecting the production of gasoline and the perfection of internal combustion engines which burn crude oils. The actual displacement of the steam by the gasoline engine puts in a new light the consequences which are sometimes associated with an estimated rapidly diminishing fuel supply.

We are, however, not concerned at present with the consequences of a condition, the facts about which are arrived at largely, if not wholly, through estimates, but rather with this method of numerically describing such condition or tendency. Attention is simply called to the fact that a very large proportion of statistical data currently collected by government and private statistical bureaus is nothing but estimates. They may be good, bad, or indifferent; but this does not now concern us. They should, however, be used as estimates, and the limitations of the methods under which they are collected fully understood. Descriptively, this method constitutes the third step in the collection process.

3. *The Collection Process* (*Functional*)

(1) Who are to be Canvassed?

Intimately connected with the statement of the purpose for which a statistical study is to be made, and the outline or plan for actual execution, is the question: Who should be canvassed? This can be answered roughly in most cases, by an inspection of the field. A complete and definite answer is possible, however, only after a directory of the possible sources of information has been completed and the types of the informants, together with the character of the material which they possess, determined by intimate study. If the problem is the fixing of a reasonable minimum wage for gainfully employed women, inquiry must be directed to those who clearly fall within the group to be benefited. If the wage is to apply to a single industry, then obviously there is a double restriction imposed. Having determined, however, the industry and the persons affected, the question remains: From whom shall information be secured? If it is secured wholly from the employer, objections may be raised that the

returns are inaccurate, that all cases are not included, that the data apply to unrepresentative seasons, that the money value of perquisites granted are included in the wages reported, that because of the stability of employment and the security of tenure, these factors are capitalized, etc. If they are secured from the workers the contention may be made that records are not kept and, therefore, that the data submitted are at best estimates, that no cognizance is taken of other things than money wages, that evidences exist that there is a strong presumption of a desire to make a case, etc. Neither source may be depended upon absolutely, yet in case of wide divergence or difference in reports or testimony, and in the absence of the actual facts, reported figures have to be taken. If any of the above considerations maintain, they, of course, may be given weight in the determination of actual conditions. A single source is not always available; frequently, it is necessary and desirable to use various sources in order to get the facts and to see them in their correct light.

If the subject of study is budgets of workingmen's families; Who shall be included and who excluded? What national, racial, customary trade, occupational, and wage boundaries to the problem shall be set up? How many budgets can be secured? How many must be taken and for what period must they apply in order to give validity and general application to conclusions? How wide must the survey be to be typical of the group or class? These questions cannot be answered off-hand; they demand careful consideration and the exercise of keen judgment and sound statistical sense.

If it is desired to test the results from the operations of a law which requires all employers of five or more persons to report industrial accidents to a central authority, and to render conditions of labor safe by the adoption of ade-

quate safety devices; Who are affected by its provisions? Failure to comply with a law cannot be made punishable when the supplying of blanks, for instance, for reporting accidents and recording the installation of safety devices, is made a condition of the law's operation, and this the administrative board has failed to do. In the administration of such laws one of the most difficult problems is the perfection and current correction of directories of those to whom the law applies. Anything like a statistical statement of the results accomplished or the conditions maintaining is impossible without determining those who are affected.

Frequently conditions of time, money, and organization require that sources of information be omitted or that typical facts alone be presented. The problem becomes one of sampling. What shall be used and what omitted? An index number of prices may materially be affected by the omission or by the too frequent use of a given commodity or set of commodities.[1] The reasonableness of a court decision or of an administrative ruling as to what constitutes a "fair return" may hinge upon the inclusion or exclusion of certain representative railroads. The omission of an important sale, under the sales method of evaluation, may materially affect the price accorded to real estate in a given district. In the determination of a unit value for urban land how much importance shall be assigned to corner influences, to frontage, and to relative position? Small deviations in either matter from the standard usually employed may make a material difference in the value assigned. The area included may be too large, conditions may not be homogeneous, and the resulting unit value not be typical. The problem is essentially one of judging the conditions to be included, together with determining the weight to be assigned to each controlling

[1] See *infra*, chapters IX and X.

factor, and is not unlike the problem of discriminating between this and that source of information, of including or eliminating this concern or that individual, in the attempt accurately to represent a group or to determine the direction of a trend.

Who shall be canvassed, and what conditions shall be included, depend in large part as to whether samples will suffice or whether all data are necessary for an adequate picture. If it is decided to employ samples only, care should be used to distribute them over as many categories as are represented in the full data, and to guard against an undue emphasis on any particular quality or feature peculiar to a given type. If one were interested in the typical wage paid to mechanics in automobile factories in the Middle Western States, obviously little weight would be given to the conditions in the Ford factories. On the other hand, if complete data on wage-rates in the industry as a whole were desired, the exclusion of the Ford Company would be a serious error.

Comparatively few workingmen's budgets, if accurately kept and reported, will serve to give a correct picture of the cost of living.[1] It is unnecessary to canvass all individuals of the class considered. The Bureau of Statistics in Massachusetts maintains that the returns from representative manufacturing establishments are superior to those which would be secured if returns from all establishments were included. What is desired, of course, is not a record of capital employed, wages paid, etc., for *all* establishments, but only for representative ones. On the other hand, in the collection of statistics of trade union membership and the amount of unemployment, it is desired to get totals for all unions. No

[1] For an interesting discussion of sampling, see *Livelihood and Poverty*, by Bowley, A. L., and Burnett-Hurst, A. R., Chapter VI, pp. 174–185 (London, 1915).

reasons exist for the employment of the sampling process — the statistics are meant to be inclusive. If they are not, the only alternative is an estimate upon the basis of the incomplete returns.

(2) The Schedule

In the preparation of schedules certain elementary principles should be observed:

1. Assurances should be given that the inquiries are made according to the provisions of law, or if voluntarily undertaken, with the hope of throwing light on some particular problem. Reasons for making the inquiries themselves, together with reasons for making them of the particular informants should either be stated or be clear by inference. Informants generally demand assurance that the law requires answers to be made, or that the purpose sought to be accomplished has some really vital end.

2. Schedules should be accompanied with stamped envelope for return.

3. Schedules should be as brief as is consistent with the purposes which they are to serve, and the questions asked should unmistakably be addressed to the problem. So far as possible, the bearing of each question should be evident from its context.

4. Units of measurements should be clearly indicated, be accurately defined, and so far as possible, conform to common usage. Definitions and explanations should so far as is convenient appear in the body rather than at the beginning or the end of schedules.

5. Rulings and columnar arrangement should be simple and definite so as to guard against the misplacing of items. In case spaces or columns are not to be used this fact should clearly be indicated.

6. Opportunities or occasions for making false or inaccurate answers should be guarded against by having the questions so far as is possible corroboratory.

7. Normally, the making of arithmetical calculations as totals, percentages, etc., should be reserved for the statistical organization, and not be intrusted to or imposed upon the informant.

8. Questions should be simple and unmistakable as to meaning, should not allow of evasive answers or of double interpretation, should not be unduly inquisitorial, should be arranged logically and in the order most convenient for the informant, should have an evident bearing on the purpose sought to be realized, should not involve duplications, should be capable of being answered by yes or no, or by number, and should always be civil and diplomatic in tone.

The sending out, returning, and editing of schedules raise some interesting problems and call for brief consideration. Normally, all the schedules should be sent out at one time. This process will often allay a suspicion which might arise in case one of a group receives his schedule far in advance of others. He may feel that he is being singled out for special inquiry. By schedules carrying announcement of the terms of the law, of the scope of the inquiry, or by being mailed simultaneously, inattention to details may best be obviated and coöperation secured. Moreover, the simple expedient of sending out schedules simultaneously tends to guarantee against their being late in returning, and against interference with the process of tabulation and analysis. If returns come straggling in over long periods it is often difficult to know when to "close" a case, and what to do in cases of exceptionally late returns. Second or subsequent requests may always be made, but the amount of pressure which may be applied in case of a failure to report will depend upon the

importance assigned to a given return, to the mandatory power possessed by the inquirer, to the degree of coöperation which maintains between the informant and the person or organization seeking the information, and to the period available for delays, or the position arrived at in the process of tabulation and analysis.

When schedules are returned, whether this is done by informants, or by representatives of the collecting agent, a certain amount of checking, editing, and revising is necessary before they can be accepted and the work of tabulation begun. If agents of the collecting unit send them in, a greater degree of uniformity of detail in their makeup will undoubtedly exist, and occasions for correspondence and personal interview regarding the meaning and force of certain entries will largely be obviated. The services of agents in these cases are employed before the entries are closed rather than after the schedules are received.

Upon receipt of schedules, evident errors due to omissions, addition, false entry, and confusion of items can readily be corrected. Undue tampering with the facts, however, is dangerous, and alterations should be made only in cases of unmistakable error. It is an easy matter materially to change the results of a canvass and to distort the truth by the interchange of a few items. The will to deceive or to make a case may not be present at all, and yet the same results follow as if it were present. If questions have been uniformly misunderstood, the basis for change is certain. If, however, the relationship between items is made to fit a predetermined order, then the data are used merely as a support to individual opinion.

The degree to which omissions may be allowed or error countenanced is also of great importance. If an entry on the samples used tends unmistakably to confirm a given fact,

and the samples are representative, then the omission of this fact on a number of schedules may be tolerated. If, however, the evidence tends to arrange itself on either side of a question in about equal proportion, and the drift of the trend or the degree of relationship is indefinite, then the omission of an item in a comparatively few cases may be a serious matter. It may be that these are the very items which are needed to decide the case in point. No rule can be formulated nor general principle stated which will cover all such cases. If the range for discrimination is wide, the final analysis may be determined by the judgment of the editing official.

Many of the same considerations apply in the case of error. If errors tend to correct each other, a considerable degree of inaccuracy may be allowed. If, however, they tend to become cumulative on either side, then their presence is of serious consequence and every effort should be made to remove them.

These considerations may be given point by relating them to a case where editing is of vital concern. In the use of the "sales method" of evaluating real estate the above considerations are of primary importance. All biased errors must first be removed. These are interpreted to include, among other things, cases of nominal consideration, transfers between relatives, sales involving land contracts or other conditions which in any way cloud the titles, etc. Only sales between ready and willing buyers and ready and willing sellers and involving full warranty deeds are held to be valid for use. By insisting upon these conditions, however, the number of sales actually available as a basis for value determination may be so few as to be inadequate. Shall sales made between relatives be included when the values represented by them essentially agree with the findings when

they are omitted? To include them would be to add weight to the value arrived at on the basis of other sales, providing the value thus determined is warranted. If it is not warranted, then their inclusion only tends to support a case which in and of itself is incorrect, and weight would normally be given to the conditions under which the sales were made. Their inclusion, on the other hand, may change materially the values assigned to a given district, and yet from every side the evidence is clear that they represent true value. The only consideration against them is the relations of the grantees and grantors — relations which will normally not be allowed in the use of sale statistics for the determination of land values. Moreover, how many sales are necessary to establish a unit value? With twenty sales the unit value is $100 per front foot; with twenty-five sales the unit value is $105, and with eighteen sales $95. How many sales should be included and to what districts should they apply?

Such considerations as these are vital, and their force is constantly being experienced in actual statistical work, no matter whether it applies to land valuation, price determination, studies of wages, cost of living, or what not. The function of the editor calls for the possession of sound judgment and the exercise of keen discrimination.

VI. Conclusion

This chapter has to do with the sources of secondary and the collection process of primary data. The aim is to discuss the practical steps to be followed in statistical work. Both are held to be anterior, but at the same time vital, to all other considerations in the statistical process. The discussion is intended primarily as a manual of instruction rather than as an encyclopedic treatment. If the points of view

here developed are kept constantly in mind, and there is real desire to profit by them, subsequent steps will be easier and the reader will have the assurance that he is employing in a scientific manner a delicate, though frequently abused, instrument of study.

The personal element stands out as an important factor in all that has been said. Statistics do not answer questions or support conclusions independently of the one who manipulates them. Judgment, candor, and integrity are necessary at every step. One must not only know the field in which he is working, its statistical possibilities, and what has been done, but he must also realize the difficulties under which data are collected, the precise manner in which they are used, the sources and possibilities of error and bias, etc., and the ways of detecting and eliminating them. In a word, he must understand what is involved in the preparation of an intellectual tool, and then in the light of his knowledge use it intelligently for the purpose in mind. If it is faulty he should know and acknowledge it. If it is well fitted for his purpose, that fact should be evident in the uses which are made of it. To be a good statistician one has to be more than a technician, but technique cannot be ignored.

References

King, W. I. — *Elements of Statistical Method*, Chs. IV, V, VI, and VII, pp. 39–64.

Bowley, A. L. — *Elements of Statistics*, pp. 23–63.

—— — *Elementary Manual of Statistics*, Ch. VIII, pp. 64–70.

United States Occupation Statistics, 13th Census, Washington, D. C. 1910. Vol. IV, pp. 17–23.

CHAPTER III

UNITS OF MEASUREMENTS IN STATISTICAL STUDIES

PASSING from the more general statement of the principles involved in the collection process, and of the methods of collecting statistical data, the significance of such expressions as units of measurements, purposes of studies, schedules, etc., will be clearer if they are discussed separately and studied in connection with concrete problems. This is done in the following two chapters.

I. THE MEANING OF STATISTICAL UNITS OF MEASUREMENTS

The statistical approach to a subject is always numerical. Things, attributes, and conditions are counted, totaled, divided, subdivided, and analyzed in this approach. We do not deal alone with single instances or with rare occurrences, but rather with aggregates.[1] The statistical process is both analytical and synthetical, and numerical considerations and preponderances of evidence are the chief bases for conclusions.

The numbers of aggregates dealt with always relate to units of measurements characteristic of the things or conditions studied. It is not 1000 as an abstract unit of frequency which is considered, but 1000 farms, industrial establish-

[1] "Statistics . . . does not deal with a single homogeneous mass but with a complex body composed of multitudinous units differing in form and action one from the other; and it is with the complex not with the units that it is concerned." Bowley, A. L., *Elements of Statistics*, p. 262.

ments, loans, mortgages, etc. Numbers as abstract units may be combined, separated and divided indefinitely because they are homogeneous, the more or less merely indicating presence or absence of a condition represented abstractly. In physical measurements we are accustomed to add, subtract, and otherwise treat numerically units of length, width, and volume as it suits our fancy or as necessity demands. This is generally done without any necessity of re-defining the units since they are homogeneous, standardized, and unvarying as respects time, place, and condition. They do not have to be adjusted to each purpose for which they are employed. A linear foot remains 12 inches, a meter 39.37 inches, an American gallon 231 cubic inches, etc., for all uses to which they apply, and they may be combined with like units and frequently converted into terms of each other without any serious inconvenience or risk of misunderstanding or confusion.

The same cannot be said of most units of measurements which are dealt with in economic statistics. Respecting such a unit as the ton-mile, while the physical measurements remain constant, in applying them to concrete problems many counter considerations are involved. While a ton is invariably a ton, and a mile a mile, all tons, except as to the one quality weight, are not the same, nor are all miles, except as respects distance, equivalent. One ton may be bulky, low-grade freight; another ton may be compact, high-grade freight. One may be the measure of a quantity of stovepipe elbows, the other of a quantity of silks. Likewise, one mile may be of easy grade in a prairie, the other of heavy grade in mountainous tunnels. The conditions necessary to the movement of one ton one mile — the ton-mile — may be wholly dissimilar in spite of the common name which is assigned to the service. Units must be referred to the condi-

tions which they describe, and since these are widely different, combinations of them should be made only with care and circumspection. The point sought to be emphasized is that in statistics while abstract units of size, dimension, and frequency are employed they are not dealt with as abstract units, but only as reflecting conditions which produce them and for purposes to which they apply.

Respecting most units, with which the student of economic statistics deals, the fixity and definiteness which characterize such a unit as the ton-mile do not hold. Abstract quantities or frequencies representing relative abundance or absence — a more or a less — are still employed, but the conditions which they measure and the purposes for which they are used are so different for each unit that a clear declaration of purpose must always precede their definition and use. The problem is not so much that of counting units describing different degrees of intensity, abundance, or absence of the same thing, as it is counting different things which have been given the same general name. An illustration will give point to this contention.

If our problem were simply to enumerate the number of manufacturing establishments in a given district, the definition of this unit would obviously be determined by the following conditions: (*a*) The meaning of manufacturing as distinct from trading, mercantile, transporting, agricultural, etc., pursuits. (*b*) The meaning of an establishment. The definitions employed will depend upon the purpose in mind in using them. If it is to learn the number of such enterprises when the criterion of individuality is ownership, one condition maintains; if the criteria are independent existence respecting the processes involved and the management over them, independence respecting housing conditions or contiguity, independence respecting relative location, etc., then other

conditions as surely maintain. In the first case the fact of ownership determines the fact of enumeration; in the other cases, respectively, independent processes through which manufactured goods pass while under one management or ownership, the fact of being contiguous or under one roof, the fact of being located in the same political or economic jurisdiction. In these cases it is not enough to maintain that an establishment is an establishment; the identity, and therefore the number to be enumerated, depends upon the criteria which are set up. The statistical process of grouping and combining is impossible unless the units enumerated are identical in the particulars chosen as a basis for enumeration.

One other example of a somewhat different type may be given in this connection. It is desired to determine the industrial accident rate in a given industry as a basis for fixing a scale of compensation for accidents. What is an accident? Obviously, the reason for compensation is personal injury with its attendant consequences, and it is the character of the injury which serves as a basis for enumeration. All injuries involving a loss of any time howsoever slight might be thought worthy of inclusion. But since compensation is the cause for the determination of the number, only those injuries should be included which occasion an *appreciable* loss of time. What is an appreciable loss of time? To an individual who experiences the loss a reasonable amount might be any time, howsoever slight. To the employer, however, who advances the compensation, and to the public who finally bear it, a period of one or two weeks might be thought to be the minimum compensable period. But many trifling accidents may occasion a far greater loss of time than a single or a few serious ones. There would be no hesitancy about counting the serious ones, yet there might be respecting the minor ones. But it is precisely the latter which can frequently most easily

be prevented, and about which we may desire information, since precautionary measures may be taken for their eradication, which involve little added cost to the employer, increased efficiency to the employee, and the gradual elimination of the occasion for compensation.

Moreover, only industrial accidents are to be compensated. Self-inflicted injuries as well as those occurring to workmen while not engaged in industrial operations, and when work done is not a proximate cause of the injury, should clearly be eliminated, when accidents are enumerated for this purpose. Moreover, is disease contracted directly as a result of the conditions of industry an accident? Surely it is an "injury," and if injury is the basis of compensation, ought not diseases of this type to be counted in determining upon a reasonable basis? If disease contracted directly as a condition of employment is counted as an industrial injury (not "accidental," but characteristic or regular), how should instances involving impairment of health, mental or physical ability, be considered? How long a period must elapse before a condition, the result of employment, ceases to be checked against such employment? What is an industrial accident for compensation purposes?

Our problem, however, relates to the *rate* of industrial accidents. Not all occupations are equally hazardous, and to refer to industries the accidents occurring irrespective of the occupations involved, is equivalent to assigning them to conditions which the latter cannot produce. Moreover, the number of accidents which occur is a function of the number of persons exposed to risks and the period of exposure — the man-hours or man-days. In using reported accidents as a basis for compensation care, therefore, must be taken to assign the results to conditions which produce them.

On the other hand, if the purpose in enumerating industrial

accidents were to measure the gross amount of time lost through mental or physical injury, obviously all accidents and all diseases directly attributable to industry should be included. If the purpose were alone to secure information as a basis for removing the conditions causing accidents, or for assigning responsibility for them as between employer or employee, machine or injured person, those which were trivial from the point of view of the individual would take equal rank with those denominated severe. What is an industrial accident?

Inquiries similar to the ones suggested above respecting accidents must always be made and answered before the collection of primary material or the use and analysis of secondary data respecting any problem is begun. It is not sufficient to study mere frequency, but frequency related to the units chosen, and the units in their particular applications to the cases under consideration. Too often we are prone to treat statistical data as though frequencies were abstract conceptions; to add, divide, and subtract them with little regard to their particular significance, and to their application and function when subjected to new and different uses. To formulate the purposes for which statistics are to be collected and used is the first step in statistical studies; rigidly and unmistakably to define the units of measurements in which the aggregates are expressed and to adhere to them throughout the process, is the second. The latter is governed by the former, as the former is determined by the latter. The two are reciprocal. Statistical units cannot be defined outside of the purpose of their employment, and the purposes cannot be outlined in detail with sufficient accuracy for execution without a clear notion of the units.

Probably enough has been said to bring to the reader's attention the problems in and the necessity of the accurate

determination of units, as preliminary to statistical investigations, as well as the distinctions between the use of abstract units of mass or frequency in mathematical calculations and the use of the same abstract concepts applied in statistical studies. Statistics is more than arithmetic. It is numerical, but its function is broader than numerical computations. It is concerned, as has been said, with the processes and methods of formulating and testing conclusions from premises resting solely upon numerical bases.

Leaving this more general discussion of the nature of statistical units, we may now address our attention to the types of units which should be distinguished, and to some of their peculiarities.

II. Types of Statistical Units of Measurements

Distinction should first be made between *units of enumeration or estimation* and *units of exposition or analysis*. The first are those which are employed in the collection or summation of primary or secondary data, — the units in which measurements are made, — while the second are those by means of which data are applied to problems. The former are primarily units of collection; the latter primarily units of analysis. One is related more to statistics as numerical facts; the other, more to statistics as methods in the use of these facts.

1. *Units of Enumeration or Estimation*

Units of this type may conveniently be divided into two classes, simple and composite. A simple unit is one in which a single condition is present which calls for definition. Examples of this type are: a farm, a ton, an accident, a strike, a lockout, an immigrant, a room, a street, a draft, a

bill of exchange, a deposit, a novel, a citizen, etc. The distinguishing thing about such units is the absence of any limiting qualifications. Many considerations must be given attention in accurately defining them, but the difficulties are significantly less than would be experienced were a limiting word or words added to each. When such limiting words are added, simple are converted into composite units. For instance, a farm as a simple unit becomes composite by adding such a limiting expression as "improved." The problem is now not only to define a farm, but also to define the condition known as improved. Similarly, the other units named above may readily be converted into the composite type. A ton becomes a freight-ton; an accident, an industrial accident; a strike, a carpenters' strike; a lockout, a building trades' lockout; an immigrant, a southern-European immigrant; a room, a sleeping room; a street, a business street; a draft, a sight draft; a bill of exchange, a finance bill of exchange; a deposit, a time deposit; a novel, a religious novel; a citizen, an "undesirable" citizen, etc. While limiting words undoubtedly restrict the field in which units may be employed and narrow the concepts materially, they clearly bring into play in each case two or more sets of conditions to be defined where formerly there was but one. Greater discrimination is required in order to fix the limits in which they are employed, and two or more occasions for error are introduced — errors respecting both the original concepts and the limiting words.

The composite type is not restricted to instances where only two sets of considerations apply. There may be more than two conditions which it is necessary to fulfill. For instance, a southern-European immigrant as a composite unit may be still further restricted by adding the words Christian and literate. The unit then becomes "a literate-Christian-southern-European immigrant," and, of course, in

this form is much more difficult of accurate determination than was the simple unit "immigrant." Each portion of the unit must be specifically defined and the grounds for distinction unmistakably set forth.

Moreover, a limiting word or words frequently change the meaning and significance of the simple units from that which they possess when used alone. For instance, the unit "room," in a survey conducted solely to determine the size of rooms in tenement buildings, would be defined in such a way as to call for the listing of any portion of a house habitually used as a place of abode set off by walls with exits either closed or capable of being closed. To add to this unit the limiting word "sleeping" suggests so many considerations respecting light, ventilation, size in respect to number of occupants, and time of occupancy, etc., as to alter materially the meaning attached to the unit when the counting was undertaken to determine size, but not size in connection with use.

In the case of composite units, whether made from primary or from secondary material, care should be used not to combine limiting conditions without first accurately determining those maintaining when they were separately employed, and the necessary effect of the combination. To repeat, statistical processes are not confined to counting or combining abstract units, but units defined under particular circumstances and addressed to particular problems. For instance, it is desired to compare the illiteracy among southern-European immigrants and the American negro in the Southern States. It would be clearly an error to make this comparison until the meanings of immigrant and negro were definitely settled, until comparable sex and age classes were specified, and until the same or comparable tests for determining illiteracy were employed. The illiteracy tests

for the immigrants may not have been comparable with those used for the negroes. The tests for the immigrants may not have been adjusted for the different age classes, nor shaped upon standards characteristic of the new world. Moreover, they may have been influenced by the standards used to distinguish immigrants from non-immigrants. It is indispensable for the student to define units of measurements for use in primary studies so as to serve specific purposes, and in using the units of secondary material to satisfy himself fully of their peculiarities before employing them for purposes of comparison.

The point which is sought to be emphasized is the necessity of reducing the conditions in every unit to a homogeneous basis. Conflicting and overlapping conditions cannot maintain. These considerations are of distinct application in the field of cost accounting where it is necessary that cost data be reduced to their most elemental units. If composite or compound units are dealt with, comparisons, except under the most favorable circumstances, — circumstances which seldom if ever exist, — are exceedingly dangerous. This connection is forcibly brought out in the following citation in relation to the use of cost units in New York City.

"An example of the weakness of the usual cost data is shown by the cost per square yard for certain paving work done by five different gangs under different foremen. I have in mind a single day's work for these gangs. The work to be done was identical yet the cost ranged from $1.11 per square yard to $1.89. This cost data was worthless on its face because it did not analyze the cost into the constituent elements. It accepted the *compound*[1] unit cost as final. By going back of the unit cost per square yard we find the reason for the difference in cost for doing the same thing under similar conditions. We base everything on *elemental*[1] cost data. By this is meant the unit cost of each element that enters

[1] Italics mine.

into the performance of a thing as, for instance, the laying of a square yard of asphalt pavement. The fact that it cost only $1.70 for laying a square yard of asphalt pavement is absolutely useless and misleading unless we know all of the facts entering into the cost of laying the pavement." (Here follows a statement of thirty elements to be considered in making such comparisons.) . . . "The fact is that one square yard of asphalt may be cheap at $2.00, while another square yard may be high priced at $1.00.

"Another trouble with *compound*[1] units cost data is that it compares entirely dissimilar things with each other. . . . The number of square yards to be done has a marked effect upon the unit cost per square yard and the conditions under which the work is done will have an even more marked effect."[2]

2. *Units of Exposition or Analysis*

The second type of units distinguished above are those used in applying primary or secondary material to problems. The feature which most clearly distinguishes this group from simple and composite units is their functional use. Comparison or the establishment of relations is always involved. The problem is to relate numerical facts to conditions producing them. Relations are established, and to the units resulting we apply the general term *coefficients*.

The group may be divided into two parts: (1) units of interpretation; (2) units of presentation. Respecting the first, three subclasses, or more properly, three aspects are distinguished, viz., those of condition, those of time, and those of place. The characteristic features of each subclass and the reasons for differentiating the concept in this manner may best be illustrated by means of examples.

[1] Italics mine.

[2] Adamson, Tilden, "The Preparation of the Estimates and the Formulation of the Budget — The New York City Method," in *The Annals of the American Academy*, November, 1915, Whole No. 151, Vol. LXII, at pp. 253–255.

(1) Units of Interpretation

By the use of clearly defined simple units of measurement, suppose the exact number of deaths from infantile paralysis occurring in a given year have been determined for a given district. The population of the same district has also been correctly enumerated or otherwise determined. The problem is to express the death rate from this cause in the form of a coefficient — to relate deaths to population. Obviously, the total population is too broad a basis, since the particular cause of death is common to only a restricted group of the total. Before a coefficient is established, the base should be narrowed so as to include only those of appropriate age groups, or the number of deaths occurring be corrected in accordance with the age composition of the population. Other examples will make clearer the importance of relating phenomena to conditions producing them. The marriage rate is properly related only to population of marriageable age; the birth rate, only to total marriages or to married population; the suicide rate by sexes, to adult population by sexes; occupational mortality, to occupational exposure for identical conditions; industrial accidents, to exposure in the occupations and industries affected; consumption of alcoholics, to consumers only; street accidents, to number exposed and the place and length of exposure, etc.

The distinction which is being emphasized is between crude and corrected coefficients. Frequently only crude rates are available. The use of such coefficients, however, is never to be preferred when it is possible to make the appropriate correction. It will be noted that the correction consists essentially in more accurately defining units and in applying each phenomenon rigidly to the conditions producing it. Where this is not done, comparisons for different periods or

for a single period for different places are extremely hazardous. The amount of error involved is almost never known, and therefore provision for it can seldom be made.

This leads directly to a consideration of coefficients where time or place is a factor of importance. Examples where these are vital will serve to make clear the emphasis desired. A comparison of the death rates from malaria for the South and North is of little real value. A comparison of sickness rates from spotted fever in the Rocky Mountain and New England States is meaningless; of the per capita kill of prairie dogs in Wyoming and Massachusetts is ridiculous; a comparison of the number of miles of steam railroads per capita or per one hundred square miles of territory for New Jersey and Utah is of little if any real significance. Why? The answer is clear; because the conditions are so widely different; the same phenomena are related to conditions wholly dissimilar or in each case of local application.

Similar considerations are of importance when comparisons are made between two widely separated periods. Comparison of the ratio of the number of bank failures to liabilities for the period before state and national regulations were inaugurated with the present time; of per capita city expenditures or debt of the 70's or 80's and the present time, are to a large degree without meaning. In the first case, regulation has made the conditions under which banking is now done non-comparable with the conditions characteristic of the earlier period; in the second case, the respective domains of public and private initiative have so changed that a consideration of the amount of expenditure divorced from the benefits accruing from it is without merit. Other examples might be cited, but these seem adequate to call attention to the danger in using coefficients for comparative purposes where material changes respecting either time or

place have occurred and where these have acted upon the conditions compared. It is the limitations here noticed which are frequently given expression in the hesitancy to compare American and European conditions, for instance, respecting wages, standards of life, transportation services,[1]

[1] The following cautions are of interest respecting the difficulties of comparing railway statistics in the United States and foreign countries: "Attention is called especially to the fact that the strict comparability of all the items throughout this bulletin is not assured, even by the greatest care in compilation. It would be an impossible task so to tabulate and adjust the railway statistics of a number of countries — differing from each other in so many respects — as to place them on a strictly comparable basis. Every attempt to present a comparison between statistics of different countries encounters practically insuperable obstacles to complete comparability. These spring from numerous differences in the classification of data, in the composition of accounts, and in the organization and character of the railway service. A few examples will illustrate the point.

"In most European countries the term 'freight', as employed in the statistics of freight tonnage and freight revenue, includes a large part of such traffic as is carried by express companies in the United States. . . A great part of such traffic is carried on fast freight trains along with what Americans designate 'package freight.' It is in most respects a part of the fast freight service, rather than an express service, as that is understood in the United States. Besides the question of expediency, is the impossibility — since both kinds of traffic are carried on the same freight trains — of determining for comparison on the train-mile basis the freight train-miles, in the American sense of the term, that would correspond to the revised tonnage and revenue statistics obtained by eliminating this sort of express traffic. By leaving this traffic in the tonnage and revenue statistics for freight, the data for each country are at least self-consistent.

"Differences in the character of the service affect the comparability of average receipts per passenger-mile and per ton-mile. In the case of the passenger service, practically all countries other than the United States and Canada offer a great variety of accommodations. And in those countries the cheaper accommodations, much inferior to that of the usual 'day coaches' here and in Canada, are far the more extensively used. As a result, the average revenue per passenger-mile is greatly reduced on account of the preponderance of traffic in the second, third, and even fourth classes. No allowance can be made for this difference by any adjustment . . .

"In the case of the freight service the railways of the United States carry freight to a far greater extent in wholesale lots than in any other country except Canada. European countries, including England, cater to frequent, quick delivery of small shipments. The result is a more expensive service and a higher average charge. Furthermore, the average length of haul in the United States is . . . greater than in any other country. A comparison of the average receipts per ton-mile from the freight traffic as a whole

state monopolies, city and state revenues and expenditures, etc.

Too great care cannot be taken to make comparisons legitimate. This is particularly true in the case of statistical comparisons, since they are numerical and seemingly exact. A numerical statement of a fact is often taken by the unwary and uninitiated, as sufficient proof of its absoluteness and finality, and is made to support predetermined conclusions or premises to which it has no relation. A rigid adherence in the collection of primary, and in the use of secondary data to the principles here formulated respecting units, will help the reader to use statistical facts in a scientific manner.

(2) Units for Presentation

Coefficients may also be regarded from the point of view of units for presentation. The thought is closely associated with Tabulation [1] but it appears more logical to consider it in this connection because of its relationship to the principles outlined in the preceding discussion.

The dominant thought here, as before, is the necessity of relating facts to the conditions producing them. The approach, however, is different in that the aim in this connection is to adopt that unit of time, place, or condition for presentation which will give the facts vitality and make them serve most fully the purposes for which they were collected or assembled. Statistics collected without a well-defined purpose are seldom of much value because of the lack of

in the United States and other countries is thus not a comparison of receipts for quite the same kind of service." "Comparative Railway Statistics, United States and Foreign Countries, 1912," *Bureau of Railway Economics, Consecutive No. 83, Miscellaneous Series No. 21, 1915*, Washington, D. C., pp. 7–8.

[1] *Classification — Tabular Presentation, infra,* Chapter V.

care in their preparation and because of the absence of a controlling purpose in their presentation.

> "Science has derived very little or no benefit from the miscellaneous collecting and grouping of facts without any previous notion of what they are likely to reveal. An investigation is usually made for the purpose of answering a definite question, or of verifying an anticipation. With some such end in view, with some principle by which the classification is guided, the result usually reveals not only what is looked for, but frequently still more fundamental characteristics. . . ." [1]

Too frequently the unit groups into which facts are assembled are so broad, purposeless, and indefinite that whatever value the facts may have had as collected, is lost by the failure to correlate the method of presentation with the purpose or function which they are to play. Thus we have death rates tabulated by districts so large that correlation of deaths with their respective causes in detail is impossible. From an administrative point of view such statistics are almost worthless. Contrariwise, the groups of causes of death as tabulated are frequently so broad and ill-defined as to make it impossible to single out from the groups the significant causes and to use the statistics as a basis for a health crusade. Again, density of population — a common coefficient — is almost worthless when assigned to so large a population and so diverse conditions as those found in cities of appreciable size.[2] Density as a coefficient is significant where overcrowding is a problem. Not all sections of cities are capable of producing the unit of density assigned to the entire district, while in many sections the density is far greater than the single unit implies. In some districts density is of no

[1] Cramer, Frank, *The Method of Darwin: A Study in Scientific Method*, p. 92.

[2] *Cf.* Bowley, A. L., *The Measurement of Social Phenomena*, pp. 40 ff.

significance; in others, it is precisely the unit which is most vital. The units for presentation should always be chosen with the thought in mind of making the statistics function.

Taking an illustration from a more strictly economic field, a large part of our wage statistics, as presented for public consumption, suffers almost beyond redemption because they are reported as undifferentiated totals, as average wages, or in groups so broad as to conceal the facts which they might otherwise reveal. It is of little significance to know that the great majority of wage earners in the United States receive less than, say, $1200 a year. What is necessary to know is the distribution and wages of those below this limit. The wages paid to a non-homogeneous class expressed as a total or as an average without classification is of little significance in throwing light on problems on which we need light, such as the distribution of wealth, a sound basis for arbitration of wage disputes, standards for minimum wages, etc. The units for expression are generally too broad; the facts are related to conditions which do not produce them. Statistics in this form becomes more an end than a means to an end, more a goal than a process.

Expense and time are frequently urged as serious barriers against detailed presentation of facts. The validity of this common excuse for inefficiency and for statistical sinning is not always of easy determination. Neither is the excuse always of equal merit nor of universal application. Frequently, respecting public bureaus, this excuse has in reality little weight because their activities are characterized by a lack of coöperation, planlessness, and duplication. In studying the output of statistics on economic topics there are frequently excellent grounds for repudiating such excuses and abundant reasons for characterizing public statistical activities as undertaken largely irrespective of costs. It is

not money and time which constitute our gravest statistical needs; it is coöperation, planning, correlation, and above all, an appreciation of the fact that statistics are far more than records, that they may serve as a record of achievement or the lack of it at the same time that they are made functioning instruments. They find their chief justification in the manner in which they minister to our economic needs.

III. Rules for the Use of Statistical Units of Measurements

Our general conclusions respecting the functions of units of measurements and the rules and cautions which it is necessary to follow in their use in statistical studies may conveniently be summarized as follows:

1. Refer all units of measurements to the conditions which produce them. Make them homogeneous, suited to the purposes for which they are employed, and use them with consistency and integrity.

2. Define all units clearly and fully in the beginning. Certain corollaries follow from this general rule:

(1) Study problems in all their aspects before defining the units. Anticipate all the difficulties there encountered, and make provision, if possible, for others not seen.

(2) Define all units in the light of the intelligence of the informants and the character of the data from which the facts are drawn.

(3) Make all definitions in such a form that exceptions will readily be detected, misunderstanding of terms difficult, and employment ready, and in terms and form characteristically employed. A "farm," for instance, as defined for statistical purposes, should be essentially the unit as commonly understood.

(4) Establish a logical basis for all definitions.

(5) Avoid substantive or descriptive units where direct ones are available. The unit, college graduates, for instance, is not equivalent to the unit educated persons, nor is the number of insane accurately reflected by the number of asylum inmates and the commitments to insane asylums.

3. Appreciate the fact that statistics should be viewed functionally; that a main source of error is in the units employed in the collecting, assembling, and interpreting processes, and that rigid adherence to the principles here developed respecting units is essential in their employment in statistical studies.

References

Bowley, A. L. — *The Nature and Purpose of the Measurement of Social Phenomena*, pp. 29–97.

—— — "The Improvement of Official Statistics," *The Journal of the Royal Statistical Society, 1908*, Vol. 71, pp. 461–469, on "The Nature and Condition of Statistical Measurement."

Zizek, Franz. — *Statistical Averages*, pp. 25–33.

Watkins, G. P., "Statistical Units," in *Quarterly Journal of Economics*, Vol. XXVI, pp. 673–702.

CHAPTER IV

PURPOSES OF A STATISTICAL STUDY OF WAG UNITS OF MEASUREMENTS, SOURCES OF DA SCHEDULE FORMS—ILLUSTRATIONS OF METH(

I. The Problem in the Study of Wages Statec

1. *Introduction*

In the preceding chapters emphasis has been placed u the logical order in statistical studies — deciding upon merits of the statistical approach, outlining fully the] poses of study, defining units, collecting primary and sembling secondary data. In the chapter immediately ceding this we have given concreteness to the difficultie defining and using statistical units, and have shown reciprocal relations between them and the purposes for wl they are used. We shall now demonstrate this relation n fully by studying the problem of wages and by relatin; it methods of securing primary and the sources of secon(data.

Much is now being written and spoken on the topi wages. Socialists are condemning the "wage" syst social workers and those interested in ameliorating condition of the poor are constantly urging the paymen a "living" or of a "minimum" wage. Wages is the bor contention in industrial disputes, and by some is thoı to be the ultimate source of all our industrial ills. Effici(advocates are studying various methods of wage payn

in an attempt to harmonize the principles of industrial efficiency with the interests of employees and thereby to enlist their support in having them adopted. Others are testing the level of wages in terms of their purchasing power either to measure their trend or to demonstrate their reasonableness. Still others are attempting to adjust to an increased nominal wage scale the prices charged for commodities and services in the hope of "making both ends meet." To the employee wages are too low; to the employer, wages are too high. To one, they are income, to the other, costs. The absolute commanding importance of the subject in all its vagaries is ample reason for choosing it in order to illustrate certain principles in statistical methods.

It has been thought best to approach the problem from the standpoint of a public bureau collecting data from many employers, rather than from the standpoint of a single employer assembling wage data in his own establishment. The first approach, in a sense, includes the second, inasmuch as each employer must organize the material in his own plant before filling out the schedule for the collecting bureau. Moreover, employers are vitally interested in the wages their competitors are paying, and the only available sources for the necessary facts are the reports which public bureaus are authorized to make. They are likewise interested in the collection process, for only by a full knowledge of it are they in a position to appreciate the limitations and the virtues which collected data possess. The finished product is the basis for any comparisons which they may desire to make and consequently its scope, its merits, and demerits are of vital interest to them.

To employers it is the comparative view of wage scales, methods of payment, etc., in wage disputes, arbitration proceedings and the like, when dealing with employees;

and the total wage bill, etc., when dealing with competitors, which are most important. The difficulties experienced in the collection of wage data, the value of comparisons which are currently made, and the legitimacy of claims which employees advance respecting wage-rates and hours can be appreciated only by a study of the data themselves. In what follows, we have attempted to describe the types of wage data collected in the United States, to indicate the sources from which they are drawn, with their relative advantages and disadvantages, and to suggest their probable value in the light of the principles of statistical methods.

There is another reason for approaching the problem in the manner followed. Units of measurements and schedules or reports are generally standardized in individual establishments. As between and sometimes within a single establishment, however, they may differ materially. For this reason comparisons are often of little value, although they are given much weight, and it is the dangers involved in making them which are here given particular attention. These dangers are traceable to inaccurately and loosely defined units of measurements, to unrepresentative, biased, and crudely tabulated data, and to the failure to perceive the limits of the statistical method and to abide by its principles. In order to use statistics with discrimination and integrity it is necessary to have a knowledge of their source, of the interpretation given to the original entries, of the groups and combinations into which they are thrown, etc. It is with these thoughts in mind that so much attention has been given to units in the preceding chapter, and that in this one the collection process for a concrete problem is discussed from beginning to end.

2. *Characteristic Confusions in the Use of the Term "Wages"*

The meaning of the term "wages" in current discussions is generally clear from the context in which it is used. When the term is employed statistically, however, its various uses frequently cause misunderstanding and confusion. Wages and earnings are often used synonymously without any seeming appreciation of their differences. Wages and wage-rates, nominal or money rates and real wages are used interchangeably, or at least without clear distinction of the differences involved and the conditions upon which they rest. The term "salaries," as contrasted to wages, is used to distinguish large and regular from small and precarious incomes, notwithstanding the fact that the bases chosen are in part illogical when income as salary is less than income as wages. Moreover, the criteria by which the two are distinguished are not standardized; the rules set up are not always strictly adhered to and statistical studies based upon current distinctions or in violation of them sometimes lead to grotesque conclusions. The principles developed in the preceding chapter of relating facts to the conditions producing them, and of making comparisons involving considerations of time, space, or condition legitimate, are constantly being violated.

The reasons for and types of confusion in the use of this expression will more clearly be seen by studying various purposes for which one would wish statistical information on wages and by defining the limits of the term as used for these purposes. No attempt is made to cover all, but only those purposes which bring out the peculiar statistical difficulties to which it is now desired to call attention.

3. *Bases for a Definition of Wages*

Wages are defined in current economic discussions "the income received on account of labor performed "the price of labor hired and employed by an *entrepreneur* or as including "all earnings assigned to men for their wc from lowest piece wages to highest annual salaries ε 'wages of management.'"[3] In a still different sense term is used to indicate "the share of the annual product national dividend which goes as a reward to labor, as (tinct from the remuneration received by capital in its vari forms."[4] The term thus defined is too indefinite for s tistical use, yet the distinctions suggest the differences which it is desired to call attention. The first sugge property as contrasted with service income,[5] but does distinguish money income from real income nor sala from wages. The distinction between the wage system ε other possible methods of service remuneration is reflec in the second, while the last calls attention to a use restric to economic theory — namely, that of distinguishing reward of labor as contrasted with the reward of landlc and capitalists.

A number of distinctions must be made in order to the term in statistical studies. Wage-rates must be tinguished from earnings; nominal rates from real rat and earnings from labor — wages — from earnings from sources including returns from investments, rents, It is necessary also to distinguish wage-rates from sal:

[1] Johnson, A. S., *Introduction to Economics*, p. 152.
[2] Gide, Chas., *Principles of Political Economy* (Second American Editi p. 487.
[3] Seager, H. R., *Principles of Economics*, p. 244.
[4] Webster, *New International Dictionary*.
[5] See Nearing, Scott, *Income*, Macmillan, 1915.

rates, and wages (wage-rates times the period for which paid), from salaries (salary-rates times the period for which paid). In converting wage-rates into wages the former must be increased by the money equivalent of concessions and perquisites and decreased by the money equivalent of time lost for which no compensation is received. Money wages must clearly be differentiated from real wages, or "the purchasing power of nominal wages measured by a constant standard." When computing real wages and making allowance for concessions, perquisites, payments in kind, and unemployment, the nominal money equivalent must be reduced to its purchasing power and added to or subtracted from, as the case demands, the money wages similarly reduced.

4. *Wages Defined*

The term "wages," therefore, will be used to suggest various concepts but always with the following meanings:

By wages, when used alone, are meant earnings in money or its equivalent because of manual, mechanical, or clerical labor service, paid according to a stipulated scale, at frequent intervals, and under conditions which make it customary to make deductions for short periods of time lost. This definition does not admit of the term being used to cover labor's "share" in contrast with the shares of capital and land in distribution.

By wage-rates are meant the predetermined rates at which manual, mechanical, or clerical labor service is remunerated. Wage-rates multiplied by the period for which paid equal wages as defined above.

By salaries are meant earnings in money or its equivalent because of responsible, supervisory, or directive labor service, paid according to a stipulated scale at infrequent in-

tervals and under conditions which make it customary not to make deductions for short periods of time lost.

By salary-rates are meant the predetermined rates at which responsible, supervisory, or directive labor service is remunerated. Salary-rates multiplied by the period for which paid equal salaries as defined above.

By earnings, when used alone, are meant money incomes or their equivalents received from labor services, without distinction between wages and salaries. The same term, in order to include other income than that regularly received from labor service, must be accompanied by a limiting expression.

By real wages are meant the equivalents of money wages in economic goods measured in terms of a constant standard of value.

Some of the purposes for which statistical studies of wages, as currently understood, may be undertaken, and the meaning which the expression must have in each case will now be discussed.

5. *Studies of Wages and the Uses of Terms*

If the purpose of study were to approximate the effect of trade unions upon wages one would be inclined at first to restrict the study to wage-rates, since minimum scales are determined by unions in bargaining with employers. Union figures on wages are invariably quoted as rates and are usually nominal and minima. The actual rate received is frequently higher than the minimum specified and in some cases may be even lower. If by wages are meant earnings from manual, mechanical, or clerical labor service, then the effect of union activities on employment would have to be considered. Wage-rates may remain the same and still wages be

materially affected. This fact introduces other difficulties. Are employment, strike, and other benefits to be considered offsets to wage losses or are they to be considered to be counterbalanced by increased dues necessary to replenish depleted unemployment, strike, and sickness funds? Union activities may seriously affect wages but have no influence on earnings from other sources. Wages, therefore, must be distinguished from earnings, if the latter are used to include earnings from other than labor services.

When "minimum" wages are discussed, undoubtedly wages are understood to mean rates, since employers are not compelled to hire labor but only to pay at least the stipulated minimum to those employed.[1] On the other hand, when the term "living" wage is used, reference is not so much to the rate of wages nor even to wages alone from labor service, but to earnings from all sources under the conditions possible for the persons affected. Undoubtedly, earnings from other sources than labor service, in the cases of those to whom the receipt of a living wage is a problem, are almost negligible, yet the term "income" is more suitable than the term "wages" to describe this condition.

In comparing wages for manual, mechanical, and clerical labor service by industries, occupation, districts, etc., it is necessary to use wage-rates instead of wages, since only the former are available on an extended scale. It is next to impossible to trace individuals from industry to industry and to approximate, with any degree of accuracy for an extended period, the extent of unemployment, the amount of

[1] The order on minimum wages in the brush-making industry in Massachusetts specifically takes account of the rates to be paid. "Assuming an average scale of 50 hours and regular employment" (a rather violent assumption) "this rate (15½¢) would yield earnings of $7.75." Quoted from "Estimates of a Living Wage for Female Workers," by Charles E. Persons, in *Publications of the American Statistical Association*, June, 1915, p. 577.

overtime worked, etc.[1] It is doubtful if anything better than classified rates are procured by statistical bureaus which ask for earnings. The rates as quoted by trade union sources are always minimum and nominal and, therefore, are of limited significance in determining the economic status of the groups concerned. Those secured from employers are for a limited period — generally a week, except in intensive studies — and are not a satisfactory measure of earnings from labor service. Wages instead of rates are necessary for this purpose. The same fact applies in studies relating wages to efficiency, to sex, to nationality, to length of service, etc. Wage-rates are the only data generally available and, of course, should be used as such.

If the determination of the trend of wages is the problem to be studied, wages may mean a number of things. Wage-rates, or earnings in the broad or in the narrow sense, may be considered. Study may extend to nominal or money wages or to real wages, and may include not only wage labor but salaried labor as well. If the trend of real wages — "the purchasing power of nominal wages measured by a constant standard" — is the object of study, rates and not earnings must be used, since it is only the former of which we are in possession, or which we may secure with reasonable accuracy on an adequate scale. Homogeneous wage groups must also be used. Moreover, a logical basis for the inclusion or exclusion of salaries must be established, care being exercised that the basis of distinction is followed throughout the entire period. Nothing is here said about the price index used in making the conversion of wage-rates into current

[1] For the difficulties involved even in an intensive study, see "Wages and Regularity of Employment in the Cloak, Suit, and Skirt Industry," etc., *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 147, June, 1914, pp. 14, 41, 42, 56.

prices or of the peculiar difficulties in adjusting the index to the classes of labor to which the comparison applies.[1]

If the purpose of a study of wages were to determine from the producers' standpoint the relative costs involved in labor service, as contrasted with rents or interest, obviously, rates of wages in the narrow sense used above would be too exclusive a category. Distinctions between salaries and wages would be necessary, since the purpose is merely to determine production costs assignable to labor as distinct from land and capital. If the approach to the same problem were made from the social viewpoint, it would be necessary to distinguish between wages and salaries, and on grounds other than those generally followed, inasmuch as those are frequently illogical and indeterminate. Merely to call one group salary receivers and another group wage receivers results in confusion when the economic conditions of both are similar, and when criteria for determining the status of one apply with equal force to the status of the other. There would be the same reasons for accurately defining salaries as for defining wages. The bases for the definitions should be factors of importance in the study in which the units are used. It is inappropriate to contend that the conditions according to which the units are defined change with each purpose and, therefore, that such units are unsuitable for statistical uses. The premise is valid, but the conclusion does not follow. Such a claim only serves to bring more forcefully to mind a fact already considered, namely, that while abstract measures of numerical frequence are employed in statistical studies, they are not used abstractly but applied to units the limits and terms of which are conditioned by the uses to which they are put.

[1] Index numbers are discussed below, Chapters IX and X.

II. The Relation of the Problem as Outlined to Statistics of Wages

The preceding discussion has served in a general way to show the necessity of accurately defining units of measurements in connection with the purposes of statistical studies, and to emphasize the necessary points of distinction in the use of such a word as "wages," but it has probably not related, with sufficient closeness, the subject to actual statistical data and suggested the problems by which one is confronted in using wage data possible of collection or currently collected. This closer relation we shall now establish by indicating the sources for primary wage data, by discussing the difficulties experienced in their collection, by describing the types of secondary data currently collected, and finally by constructing wage schedules to be used in connection with a concrete problem.

1. *Sources for Primary Data in Wage Studies*

(1) Primary Data Directly Applicable to Studies of Wages

Primary data in the study of wages may emanate from four sources. Those secured from employees, from employers, and from union officials are directly applicable; while those from institutions such as banks, building and loan associations, insurance companies, lodges, etc., are only indirectly applicable.

a. Data from Employees

Data on wage-rates; hours of work (nominal and actual); the amount of unemployment by cause; the methods and frequency of wage payment; earnings from labor and from other sources; perquisites in the forms of bonuses, benefits,

profits; penalties, fines, forfeits, union dues; budgetary expenditures, and facts relating to age, sex, nationality, occupation, training, length of service, previous wages, etc., may be secured in whole or in part, satisfactorily or unsatisfactorily, from individual employees, in proportion as informants are wise or ignorant, truthful or deceitful, willing or unwilling to aid, and in proportion as the statistical organization used is well or ill adapted for the purpose in mind. It is impossible to summarize in a single sentence the success attainable in securing data on wages or on any other topic directly from individuals involved. Frequently, the costs are prohibitive; in other cases, where cost is not an insuperable barrier, the types of individuals dealt with and the character of the information desired make this approach impossible. The generalization, however, is hazarded that data collected from a source where personal supervision or intimate checking are impossible are likely to possess serious limitations respecting all topics which in any way call for discrimination, for the exercise of judgment or the use of records, etc., on the part of the informant, or in which the personal equation enters to an appreciable extent. The discussion, in Chapter II, of *The Collection Process* is particularly applicable in this connection.

b. Data from Employers

Much the same types of wage data as those listed above are theoretically obtainable from employers, and the chances are much greater that they will be free from error since less ignorant groups, recorded facts, impersonal relations, etc., are dealt with. The facts, however, are of a somewhat different sort and rarely apply to an extended period. The best that can be done in most cases is to secure cross-section

views at widely separated intervals. Moreover, for the most part, classes and not individuals are considered. These may or may not be homogeneous, and in this respect are much less desirable statistical units than are individuals. From this source, with an adequate statistical organization, and with sufficient sanction, the total wage bill, time- and piece-rates, by occupations and processes, classified wage-rates, perquisites allowed and penalties assessed, and the number of employees classified by sex, age, and time of employment, etc., are theoretically available. The facts regularly secured on an extended scale and available for use are discussed below.

c. Data from Trade and Labor Unions

In many respects the records of trade and labor unions are satisfactory sources for wage data. Theoretically, nominal time- and piece-rates for regular, for overtime, and for Sunday and holiday labor; nominal hours per day and per week; benefits allowed, classified by the amounts paid, by purposes, by duration, etc.; union dues; numbers unemployed, classified as to causes, and wage losses, etc., — are available from this source. The data, however, may have serious limitations. Frequently, the desire to make out a case is held to be sufficient cause for furnishing defective returns or for withholding information. In many instances the inquiries addressed to union officials concern matters about which they can have but the most inadequate and superficial knowledge, and yet they are urged to give positive, negative, or numerical answers with few or no opportunities being offered for explanations. In some instances, undoubtedly, sincere efforts are made to state the truth as nearly as it can be determined; in other instances, no such

care is exercised. The value which data from this source possess is to a large degree determined by the scrutiny to which they are subjected by collecting agents.

The limitations, however, are not always to be attributed to errors in reporting or to incomplete returns. Frequently, they result from misusing and assigning finality to figures at best but estimates, from ignoring the specific advice of collecting agents, and from violating the fundamental principles of statistical methods. The same result, however, may occur respecting data drawn from the most acceptable sources. Statistical facts will be cited to prove contentions with which they have no connection and will be distorted and misapplied so long as people have hobbies, lack integrity, or are ignorant of the functions, limitations, and purposes of statistical data and legitimate ways of using them.

It will be noted that data on wages from unions are restricted to nominal rates and to union members. These are serious limitations where wages or earnings are sought and where non-union labor is involved. Such data are of little value in discussions of minimum wages, living wages, or other topics in which light is desired primarily concerning unskilled labor.

(2) Data Indirectly Applicable to Studies of Wages

Facts which contribute indirectly to a knowledge of wages and wage conditions may be gleaned from a study of the increase or decrease of savings, the number of depositors in savings institutions and the average deposit, the size of employers' payrolls, the activities of building and loan associations, the growth or decline of fraternal insurance, the increase or decrease of union membership, etc. In most respects their connection with the topic is remote and con-

tingent. They are at best suggestive and corroboratory and should be used with extreme caution, cognizance being taken of the round-aboutness of their application, the potency of other contributing causes to produce the effects shown, the interrelation of economic phenomena, etc.

Having sketched the types of wage data theoretically available, their sources, and the difficulties in securing and the dangers in using them, we may now briefly enumerate the types currently collected with their sources and some of their peculiarities. No attempt is made to describe or criticize fully or even to enumerate all forms regularly and irregularly collected in the United States. This has been done in a general way by others.[1] Moreover, such a treatment is not germane to our immediate purpose.

2. *Types of Secondary Wage Data*

Secondary data on wages collected from the chief primary sources are available in many forms. They appear in public and private reports, issued on the basis of data furnished by wage earners, employers, and unions. Some reports appear regularly, some irregularly; some are restricted to the single topic, while others bear upon it only indirectly. Some are monographs on special topics, while others are exhaustive independent surveys.

(1) Secondary Data Directly Applicable to Studies of Wages

a. Data from Employees

Wage studies, in which the material is drawn from individuals alone, are made primarily in connection with cost

[1] Nearing, Scott, *Income*, Chapter II, pp. 18–52, New York, 1915; Streightoff, F. H., *The Distribution of Incomes in the U. S.*, Columbia University Studies, Vol. LII, No. 2, 1912.

of living studies, such as those of Chapin[1] and Mrs. More[2] in America; Rowntree[3] and Booth[4] in England; or as a condition of the administration of labor laws, such as those on compensation for industrial accidents. Those of the first type generally apply to limited territories and restricted groups, cover only a relatively short period, and are made in connection with or are designed to throw light upon budgetary matters. In those of the second group, wage data are subsidiary to the main purpose of study, are restricted to definite classes, are not collected simultaneously for all groups, in some instances are semi-confidential, and are generally too meager to be conclusive respecting either ruling wage-rates or wages. Hence, they are not generally published except in summary form along with accident and other data.[5] They are, however, of excellent quality, because of the purposes for which collected, and in the course of time when they have been sufficiently accumulated will undoubtedly furnish material for thorough and comprehensive wage studies.

Studies on wages from material drawn directly from employees are published only at irregular intervals and cannot wholly be relied upon for current information. Those associated with budgetary matters refer invariably to wages or to earnings; those arising out of the administration of labor laws always relate to rates of wages. Those of the

[1] Chapin, Robert C., *The Standard of Living Among Workingmen's Families in New York City*, Charities Publication Committee, New York, 1909.

[2] More, L. B., *Wage Earners' Budgets*, New York, 1907.

[3] Rowntree, B. Seebohm, *Poverty; A Study of Town Life*, London, 1906.

[4] Booth, Charles, *Life and Labor of the People*, London, 1891.

[5] The brief tables on wages in "First Annual Report of the Industrial Accident Board," *Massachusetts Industrial Accident Board*, Boston, 1914, and in "Report No. 4" on "Industrial Accidents in Ohio, January 1 to June 30, 1914," by *The Industrial Commission of Ohio*, Columbus, Ohio, 1915, are illustrative.

first class are important in calling attention to low wages in certain industries, in certain districts, for limited groups, and are indispensable in the determination of minimum and living wage standards, but are inadequate for comparing wages by industries, by localities, and over long periods. Neither do they furnish material for measuring the trend of wages. Those of the second class may be used to correlate wage losses and amounts of compensation for accidents, but at present are in the main superficial and restricted studies, serving no other purpose than that of a record of wage data collected on accident schedules.

b. Data from Employers

The statistical matter relating to wages and wage conditions reported and published by regularly constituted statistical bureaus, by special commissions, and by individual investigators, may be divided into two groups; those directly related and those remotely connected with the topic.

(*a*) Material Directly Related to Wages

Direct material relates, first, to the total wage bill paid, and second, to classified wage-rates. The United States Bureau of the Census publishes at decennial and at certain intercensual periods the total salary and wage payments made during the year to which the census applies, to salaried officers, to superintendents and managers, to clerks, stenographers, and other salaried employees, and to wage earners including piece workers in manufacturing and mining industries. The Interstate Commerce Commission regularly publishes in *Statistics of Railways in the United States* the amount of compensation by years and the average daily compensation received by railroad employees classified into

eighteen groups, by classes of roads and by transportation districts. The same commission publishes for express companies the wages and salaries of employees in the "traffic," "transportation," and "general expense" divisions. A few state bureaus of statistics and labor, particularly those in Massachusetts, New Jersey, and Ohio, collect and publish, as part of their manufacturing censuses, the total compensation for labor services classified as salaries and wages. The schedule[1] used by New Jersey calls for the "total amount in wages paid during the year," and instructs informants that "only wages paid to wage earners actually employed" in an establishment or in "erecting or placing its products elsewhere" should be included. Salaries of managers, bookkeepers, salesmen, etc., should be omitted. The schedule[2] to manufacturers used by Massachusetts asks for the "total wages (paid during the year to wage earners only)," and instructs the informants to omit "salaries of agents, managers, bookkeepers, clerks, salesmen, and others of this class." The schedule[3] used by Ohio contains essentially the same questions and provides for the same omissions, except that salespeople are divided into two groups, traveling and non-traveling.

Classified weekly wage-rates are collected and published for manufacturing enterprises in a number of states, but most satisfactorily in Massachusetts, New Jersey, and Ohio. In those instances the data are taken from payrolls. Massachusetts and Ohio in their schedules ask specifically for weekly rates, while New Jersey apparently desires weekly earnings.[4] Massachusetts and New Jersey supplement their schedules

[1] *Bureau of Statistics of Labor and Industries.*

[2] *The Bureau of Statistics, Division of Manufacturers.*

[3] *The Industrial Commission.* It is not quite correct to speak of a "Manufacturing" census in the case of Ohio.

[4] The data are published as "earnings" but undoubtedly are rates.

by field agents. Ohio is able to dispense with these in connection with her wage studies, inasmuch as in the administration of her compensation law, she secures the audited payrolls of all employers subject to the law. It is not likely, under these conditions, that employers affected by the law in both respects will furnish incorrect returns. The schedule of classification in Ohio is by one dollar groups above $3 and less than $10. The remainder is as follows: $10 to $12, $12 to $15, $15 to $20, $20 to $25, $25 to $35, $35 to $50, $50 to $75, and $75 and over. The Massachusetts schedule proceeds by one dollar groups from $3 to $16, two dollars groups from $16 to $22, and the balance as follows: $22 to $25, and $25 and over. The New Jersey schedule provides for one dollar groups from $3 to $10, and the remainder as follows: $10 to $12, $12 to $15, $15 to $20, $20 to $25 and over. The sexes are distinguished for adults in Massachusetts and New Jersey, the age of distinction for adults and children being 18 in the former and 16 in the latter. Ohio distinguishes between adults and young persons, making the division at 18 years of age, and further classifies both groups by sex. Moreover, the classified scale in the case of Ohio extends to clerks (not salespeople); bookkeepers, stenographers; to salespeople (not traveling) and to traveling salespeople. In the other states mentioned the classified scale applies only to wage earners as they define the term. In each case the week for which the data are secured is that in which the largest number is employed during the year.

The most exhaustive study of classified wage-rates for the United States is that on *Employees and Wages* made by the Census Bureau in 1903 under the direction of Professor Davis R. Dewey, and known as the "Dewey Report." The data refer to the years 1890 and 1900, apply to thirty-three industries, but include only a limited number of establish-

ments in each industry. Wages of 103,453 employees in 1890, and of 160,859 in 1900 were tabulated in detailed groups. While the study is exhaustive in scope and unique in method it is not of current interest and must be passed over with brief mention.

The United States Bureau of Labor Statistics publishes from time to time special studies on wages and hours in different industries. Those on Cotton, Woolen, and Silk, 1907–1913;[1] and on Boot and Shoe, Hosiery, and Knit Goods, 1890–1912,[2] are illustrative. The data are for one payroll each year, apply to identical establishments, and give the average rates of wages per hour, the computed average full time weekly earnings, and the number of employees receiving classified wage-rates per hour by occupations for the sexes separately and by geographical districts. To facilitate comparisons, relative or index numbers, based on the year 1913, for the average rates of wages per hour and for full time weekly earnings, are also computed. From 1890 to 1907 the same bureau published a general index number of rates of wages per hour based upon the average wage 1890 to 1899. From 1907 to 1914 an index was computed only for those industries for which special wage studies were made. In 1914 such a study was made general but applied to union labor only.

(b) Material Indirectly Related to Wages

The material indirectly bearing upon wages may be classified under two heads, first, actual or average number of employees by months, and second, the time which plants operate during the year.

[1] *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 150, Washington, D. C., 1914.

[2] *Ibid.*, Whole Number 104, Washington, D. C., 1913.

The United States Bureau of the Census publishes for manufacturing and mining industries the number of wage earners, including piece-workers, as per payrolls or time records, on the fifteenth day of each month for the periods covered by its reports. No distinctions are made for age and sex classes. New Jersey, as a part of her manufacturing census, publishes the "number of persons employed"[1] during each month of the year for which study is made, classified by sex for those sixteen years of age and over, but without sex classification for children under sixteen. Massachusetts publishes the average[2] number of wage earners during each month for males and females separately but without age classification. She likewise publishes the number of wage earners eighteen years of age and over and under eighteen years of age classified by sex on the thirteenth[3] day of December as per payroll. Ohio requires employers to report the number of wage earners employed on the fifteenth day of each month as per payroll, classified by sex but not by age.

Ohio, likewise, requires employers to report the number of full days that plants are in operation and idle during the year, the former including part-time days reduced to a full-time basis and the latter not including Sundays and holidays unless plants normally operate on these days. The number

[1] Neither the instructions to informants nor the schedules define this number. Whether it is to be the average force computed on the basis of twenty-six, thirty, or thirty-one days, to be the normal force during the period, or the number of separate individuals to whom employment was given during each month, we are not told. It conceivably might be any one of them, carefully computed, but more likely it is a rough average representing nothing better than an estimate.

[2] The use of an average in this case seems unnecessary and to somewhat lessen the value of the figures in computing the deviations from month to month, with the purpose of throwing light on the seasonal character of employment. There seems no sufficient reason why the exact number, as required by Ohio, and others, should not be called for.

[3] This is the date indicated in the schedule for 1913.

of hours normally worked per full day or shift and per full week is also required to be reported. In Massachusetts the number of days in operation and idle is included in the manufacturing schedule and published in this form. Informants are specifically reminded that the working year is composed of a stated number of days and that the sum of the days reported, not counting Sundays and holidays, should total to this number. In New Jersey, data are published for manufacturing establishments on the number of days in operation, the normal number of hours per day, the normal number of hours per week, and the total number of hours extra time during the year in which establishments operate. The Bureau of the Census publishes like figures on the number of days manufacturing and mining establishments are in operation during the year and the number of hours normally worked by wage earners per shift and per week. Respecting the latter topic informants are instructed that "all that is desired to know is the practice generally prevailing in respect to the hours of labor of employes."

c. Data from Trade and Labor Unions

The wage data regularly collected from union sources by statistical bureaus refer to nominal (minimum) time- and piece-rates, nominal (maximum) hours per day or per week, causes and extent of unemployment, number and duration of strikes, etc. In this descriptive part of the chapter it will suffice, in view of what has been said above, briefly to describe the statistical activities of the United States Bureau of Labor Statistics, of the Department of Labor of the State of New York, and of the Bureau of Statistics of Massachusetts, respecting union wage conditions.

The United States Bureau of Labor Statistics has pub-

lished the union scales of wages and hours of labor for the principal mechanical trades, for the largest cities of the United States for the period 1907 to date. The report for 1913 covers the forty industrial cities located in thirty-two states for which the Bureau publishes retail price statistics. Union scales for both wage-rates and weekly hours are followed, but such scales fix the limits in only one direction. Minimum wage-rates are established below which members of unions will not as a rule work, and maximum hours beyond which they will not work at regular rates of pay. In certain cities and trades workmen are paid more than the union scale and work regularly less than the scale of hours. However, the Bureau takes no cognizance of these conditions. All wage-rates are reduced to an hourly basis, and for all the trades for which the Bureau has figures, relative or index numbers are computed for both wage-rates and hours for the years 1907 to 1913. The data are collected by special agents in personal visits to union business agents and secretaries, and wage scales, written agreements, and trade union records consulted wherever available.[1]

Statistics of unions and their membership were first collected by New York State in 1894 and 1895. Since 1897 such statistics have been regularly published. Information is now collected semi-annually from all unions, in part by schedule and in part by field agents. Schedules relate to membership and idleness, to hours of work, to new trade agreements, to changes in the rates of wages, and to rates of wages of time workers. The amount of unemployment is reported under six specific and one miscellaneous head; lack of work, lack of material, the weather, strikes or lock-

[1] A similar study, in coöperation with the United States Bureau of Labor Statistics, is made by the Industrial Commission of Ohio and applies to all the larger cities in the state.

outs, sickness or accident, old age, and miscellaneous. The data apply to the sexes separately and to the end of March and September as the case might be. The regular hours of work for Saturday, Sunday, and other days, and the total per week by branches of trades and for the sexes separately are included. Changes in hours, with those before and after each change, and the number of persons affected are also requested. Respecting rates of wages information is secured on the rates before and after changes, the number of members affected, and the estimated weekly earnings before and after changes in the case of piece workers. Schedules respecting wage-rates of time workers relate to each branch or grade of work, to the working hours per day for the specified rates, and to the number of members by sex receiving them. Other inquiries of less significance and certain modifications of these are also included.

The schedule is a model in technique; the questions are vital, clearly stated, and well arranged. It is mailed to union secretaries, ten days are given for answering, and delinquents are visited by field agents of the Bureau. Approximately 50 per cent of the schedules are sent in by mail and 50 per cent "fielded."

The published material is issued in two series: one called "Series on Unemployment" and the other "Series on Labor Organization." The first shows the amount of unemployment by cause, by months, and includes summaries for years by industries and by detailed trade groups. The issuance of a letter on the state of the labor market based upon monthly returns from the larger unions is also a regular feature of the Bureau's activities. The second series relate to the number and membership of unions classified so as to show data by industries, by trades, by localities, etc.

This account of the New York Bureau's activities respect-

ing union wages and conditions, although brief and sketchy, is probably adequate to reveal in a general way the types of data collected and the manner of securing them. Neither the schedules nor the methods of tabulation are open to severe criticism. The only criticism which might be offered is that the facts are supplied by unions. Essentially the same facts, but in a different form, respecting wages, hours, and unemployment, are available from employers and the probabilities are that they are more accurate when so returned than are those furnished by unions in spite of the care exercised to correct the errors. Employers are subject to state supervision in many respects, the statistical machinery is adjusted to this source of information, and the reporting of facts may be required legally. Unions are not compelled to report nor are they punished for withholding or distorting the matter supplied. In one respect, however, it seems necessary to deal with unions as units. Public and private boards of arbitration require *union* scales of wage-rates and hours as bases for making awards. These facts for *unions* cannot be gotten from employers; their scales do not necessarily express union experience. Unions must supply the material.

The Massachusetts Bureau of Statistics in its Labor Division collects and publishes statistics of organized labor relating to union scales of wages and hours, number and membership of unions, unemployment, strikes and lockouts, etc. Each of these will be touched upon briefly inasmuch as they probably represent the most accurate and complete data on organized labor now regularly collected by any statistical state bureau in the United States.

A report on union scales of wages and hours is regularly issued. The data are furnished entirely by unions and are published as reported, no inquiry being made as to the

extent to which the union scales prevail in the various trades and localities. That is, minimum rates and not those actually received by union labor are published. The process of collection may be indicated by reference to the 1913 report. Returns by schedule were received from 1093 unions, or 78 per cent of those in the state. By the use of special agents 200 more were obtained so that 92 per cent of the locals in the state were included. In tabulated form they show rates of wages by the hour, day, week, overtime (hour), and Sunday and holiday (hour); and hours of labor, by the day, week, and the period in which half-holidays are in effect, all classified for occupations and for municipalities.

Statistics on the number and membership of unions have been systematically collected and published since 1908. The collection is mainly by schedule and includes national and international unions with affiliated locals in Massachusetts, their relationship to the American Federation of Labor, the number of chartered local unions and the proportion in Massachusetts with their membership, classified for the sexes separately, by municipalities, occupations, industries, etc.

Statistics on unemployment among organized wage earners are issued quarterly. The data are collected from unions solely by schedule and are published so as to reveal the amount of unemployment by cities and occupations due to lack of work or material, unfavorable weather, strikes or lock-outs, sickness, accident or old age, and other reasons, the latter specified in detail. Approximately 75 per cent of the locals are included in each quarterly report.

Statistics on strikes and lock-outs have been collected by the Massachusetts Bureau since 1881. Unions and employers are scheduled on the basis of information supplied by newspapers, trade journals, etc. Besides certain pre-

liminary data the following facts are secured from unions: the names of employers affected, conditions demanded by strikers, conditions before and granted after strikes, who ordered strikes, the occupations and numbers of strikers (the latter by sex), the dates on which strikers left and resumed work and on which strikes were ended, as well as the methods of settlement. From employers those questions of the above which apply and the following are asked: the number of employees who struck, classified by sex; the number of non-strikers thrown out of work, classified by sex; the time lost by non-strikers; measures used by strikers to regain their positions, etc. In approximately 50 per cent of the cases the returns from the two sources are so contradictory as to necessitate the use of special agents to obtain the facts.[1] Even by this method in many cases the facts prove to be so indeterminate that the reports are published only on the basis of what *seem* to be the facts after all evidences are given their appropriate weight. These reports, therefore, appear to be summaries of reported or estimated facts concerning industrial disputes — knowledge of which is received through the press, by hearsay or by other means — having little value alone in connection with wage studies, and chiefly of interest for informational and not for functional use.[2]

Without citing further detail of the practices and experi-

[1] Estimated for the writer by the Division Chief. New Jersey, placing complete reliance in newspaper clippings for initial information and depending altogether for the facts secured on schedules from unions alone, publishes an annual report on strikes and lock-outs. If the experience of Massachusetts respecting like data is worth anything, statistics thus collected stand condemned.

[2] A detailed estimate of the value of these and like data compiled by the Bureau is not attempted here. It was made, however, by the writer during the summer of 1914 for the *United States Commission on Industrial Relations*.

ences of American statistical bureaus in securing wage and allied data from trade unions, sufficient has been said to indicate the problems and possibilities in this approach to the study of wages. In all cases nominal and minimum rates are involved and these are reported under conditions which make it difficult, if not impossible, to apply them to unemployment data in any attempt to approximate earnings from labor service. When properly checked by scrutinizing trade agreements, nominal hours and time-rates from this source may be determined with reasonable accuracy. Any attempt, however, to secure piece-rates on an extended scale from this source is bound to prove unsuccessful. Unemployment data from unions at best are approximations, and, of course, refer only to union labor. They serve fairly well to give a general notion of seasonal displacement of labor and of trade depression or boom but are of little value in measuring earnings or economic distress. Statistics of strikes and lock-outs as collected may serve as a rough measure of the frequency of labor disturbances but not of their consequences nor of the correction which it is necessary to make from this cause when estimating wages from wage-rates.

In summary, we may briefly relate the statistical data extant on wages to the various concepts which this term suggests.

Comprehensive data on wages as defined above do not exist in the United States.[1] For annual reports for all manufacturing industries on classified wage-rates for short pay-periods, where conceivably wage-rates are equivalent to

[1] Nothing is said about our present national income tax statistics. The exemption allowed is so high as to omit most "wage earners," and the returns are not published in a form suitable for estimating earnings for such groups. See Falkner, R. P., "Income Tax Statistics," *Publications of the American Statistical Association*, June, 1915, pp. 521–549.

earnings — assuming neither over-time nor time lost — we may turn to Massachusetts, to New Jersey ("earnings" in this state), and to Ohio.[1] Studies of classified wage-rates for special industries are periodically made by the United States Bureau of Labor Statistics. In order to use nominal and minimum wage-rates as equivalent to wages it is necessary to assume that nominal conditions are actual, that figures are reported accurately, and to correct rates by figures on unemployment supplied by unions, by employers, or by employees. The reliance which can be placed in union figures on strikes and other causes of unemployment has been suggested above. The importance to be assigned to fluctuations in the employed force, as indicated by the average or actual number of employees at various times in each year, depends largely upon the fluidity of labor, the ability of wage earners to find employment, and the complementary character of industries, studies of which on a significant scale have not been made. The fact of unemployment is known but it is next to impossible, except in intensive studies, to measure it by applying to those affected. The United States Census Bureau attempts to measure it from this source but the best that is secured is a rough approximation.[2] Moreover, it is chiefly among unskilled labor that unemployment is greatest, and union figures do not furnish the desired facts. Wages, therefore, in the sense in which the term is used here are not available in any other form than as estimates.

On the other hand, wage-rates for short periods, taken from employers' payrolls for manufacturing and some other

[1] Not restricted to manufacturing industries in this state.

[2] A question on unemployment was first included in the population schedule by the United States Census in 1880. The information secured, however, was never published. In the three succeeding censuses a similar inquiry was included, the form in 1910 being "whether out of work on April 15, 1910" and "number of weeks out of work during the year 1909."

industries, are reported with reasonable accuracy to a few state bureaus. In these cases, industries constitute the units, individuals and occupations being lost sight of in the grouping process. To supplement such data there are the nominal wage-rates reported by unions in which distinctions are made for occupations, industries, sexes, etc. The data are supplementary but not comparable. At least no comparisons of rates are currently published by bureaus to which both sets of facts are reported.

Earnings, in the sense of income from labor service without distinction being drawn between wages and salaries, and in contrast to property income, may roughly be approximated from the income and expenditure accounts of industrial and other businesses.[1] Our income tax returns do not aid us in this respect since we, unlike most European countries, neither distinguish between "earned" and "unearned" incomes in fixing rates nor differentiate incomes by sources in publishing returns.

III. A Study of Wages: Declaration of Purpose, Definition of Units, Schedule Forms

Without considering the types and sources of data on salaries and salary-rates, and without treating prices in relation to wages and wage-rates, we pass immediately, in order to illustrate the preceding treatment, to a discussion of a wage problem upon which it is intended to collect primary data. Criticism of the substance, form of tabulation, and interpretation of existing secondary data must rest with the brief sketch given above. The immediate problem, then, is to state definitely the purposes of the study which is

[1] See the studies of Nearing, *op. cit.*, pp. 18–52; Streightoff, F. H., *op. cit.*, pp. 44, *passim.*

intended to be made, to outline the plan to be followed, to define the units to be used, to formulate schedules and to outline suggestions for the receipt and editing of returns. The precise use which will be made of the data will, of course, be determined in part by the character of the replies and can only tentatively be outlined in advance. It is intended, however, to establish certain relations and make certain comparisons between the facts reported, and the tabulations will be adjusted to these ends.

1. *Declaration of Purposes*

The problem which has been chosen for study is the wage conditions in the textile industry in North Carolina for the year 1914. For convenience, the survey is restricted to manufacturers of cotton goods including small wares. On the basis of information collected, schedules will be sent to 100 establishments which were found to be doing this business at some time during the year, the basis for listing establishments separately being that outlined in the schedules. We are interested to know the level of wage-rates for the sexes separately, for adults and young persons, to measure the fluctuations and seasonal character of employment and their relations to wage-rates, to determine the wage bills to employers during the period, to study the relation of wage-rates to character of business organization, to fluctuations in employment, etc. The schedule is formulated with these points in mind and is intended to be filled in by employers without supervision, other than that which is received from the instructions contained in the schedules. The study is undertaken under the assumption that it has sufficient sanction, that the filing of the returns is obligatory, that returns for individual establishments are not to be published

separately, and that the results of the study will be of general social interest in which informants share equally with others. Sufficient time is to be allowed for full reports to be made and tabulations and analysis are not to be begun until satisfactory reports are received from all concerns scheduled. No attempt is made to supplement the data collected from employers by scheduling either individual employees or unions. Complementary material may be secured from those sources but in this study it is intended to rely wholly upon the returns from employers.

It must clearly be kept in mind that the discussion immediately above is illustrative of the steps which would have to be taken in the study of such a subject as wages. The facts have been given somewhat more in detail than would have been necessary had the purpose been merely to describe the data on wages and wage conditions in the United States. Moreover, it must be remembered that the requirement that all of the schedules must be returned is rather more severe than would be made in actual statistical work. The aim has been to duplicate as nearly as possible the steps to be taken in an actual investigation. Of course, it is not possible entirely to do this, but the nearer it can be done, the more interest the student will have in his work and the more value he will get from it. That which is sometimes considered to be meaningless, routine clerical work may, by paralleling as nearly as can be a real problem, frequently be thought to be both necessary and vital. Great value comes from having a student see a problem as a whole and the correlation of the different parts. By so doing the meaning of all the statistical steps through which he is led takes on new light. He is then not so much studying *method* as a problem to which method is vital in its explanation. Most mature minds desire to see some goal to their activities

and reasons for the methods of study which are used. And this is as it should be, for then individuality is bound to reveal itself and the use of statistics becomes more than mere routine labor.

2. *Schedule and Explanations*

THE X. Y. COMMISSION OF NORTH CAROLINA

RALEIGH, NORTH CAROLINA

It is desired to make a study of the wages and wage conditions for the calendar year 1914 in the establishments in North Carolina which manufacture cotton goods, including small wares. All concerns in the state doing such business are included in this survey. The study is undertaken in accordance with the provisions of law, (see Chapter 673, laws 1914) and your coöperation in making it a success is respectfully solicited. Individual returns will not be published separately, and every care will be taken to hold the facts reported confidential. All employers submitting the reports called for will be furnished gratis with copies of the complete report as soon as published.

Read the whole schedule through before answering the individual questions. *Accurate* answers according to permanent records are required on *all* questions.

Use the enclosed self-addressed and stamped envelope for returning the schedule. Schedule should be returned not later than April 30, 1915.

THE X. Y. COMMISSION,
Raleigh, North Carolina.

I hereby affirm that the accompanying report is accurate and complete to the best of my knowledge, and is made according to the permanent records of this establishment.

....	....
Name of Concern	Name of Secretary or other person making the return

....	....	....
P. O. Address	Month	Year

Schedule to be Used in the Collection of Wage Data by Establishments in the Manufacture of Cotton Goods, Including Small Wares, North Carolina, Year 1914.

1. Name of Establishment..................................

Use a separate schedule for each establishment. By an establishment is meant a plant or mill as understood in general usage. Where separate plants are owned in common, are contiguous and carried on under one set of books, such separate plants are reported together as one establishment.

2. Name of Corporation, Firm, or Individual Owner.............
....

3. Location of Factory:
County .... City or Town.............
Street and No............ P. O.. ....

The location should be that of the physical plant and not of the financially controlling head.

4. Character of Business Organization (..........) Individual (..........) Firm (..........) Corporation

Indicate whether individual, firm or corporation by checking thus (√) the appropriate term.

5. Frequency of Payment (.............) Weekly (............) Fortnightly. Time- or Piece-Rates (............) Time (............) Piece

Indicate the frequency of payment, and whether time- or piece-rates *prevail* by checking thus (√) the appropriate terms.

6. Character of Industry..................................
Indicate by giving principal product manufactured.
Please be specific respecting the principal product. The data are necessary for accurately editing the returns.

7. Number and sex of Wage Earners, both time- and piece-workers; *not* salaried employees.

Wage earners are persons receiving money or its equivalent because of manual, mechanical, or clerical labor service, paid according to a stipulated scale at frequent intervals, and under conditions which make it customary to make deductions for short periods of time lost. These should be *included*.

By salaried employees are meant persons receiving money or the equivalent because of responsible, supervisory, or directive labor service, paid according to a stipulated scale at infrequent intervals and under conditions where it is not the custom to make deductions for short periods lost. These should be *omitted*.

Age and Sex of Employees	A Greatest Number Employed at Any Time During the Year	B Least Number Employed at Any Time During the Year	C Total Amount Paid in Wages During the Year
Men 18 years of age and over .	—	—	—
Women 18 years of age and over	—	—	—
Young persons under 18 years of age			
Boys	—	—	—
Girls	—	—	—

8. Number and sex of Wage Earners employed on the 15th of each month, 1914. If data are not obtainable for this day enter the same for the nearest representative day.

Data to be of the 15th of the Month	Number of Wage Earners Both Time- and Piece-workers Employed on the 15th Day of Each Month			
	Adults 18 Years and Over		Young Persons Under 18 Years	
	Males	Females	Males	Females
January	—	—	—	—
February	—	—	—	—
March	—	—	—	—
April	—	—	—	—
May	—	—	—	—
June	—	—	—	—
July	—	—	—	—
August	—	—	—	—
September	—	—	—	—
October	—	—	—	—
November	—	—	—	—
December	—	—	—	—

9. Classified Weekly Wage-rates for the Week of the Greatest Employment during the year 1914.

Do not include over-time; short-time earnings should be reduced to a full-time basis; bonuses and premiums, if any, should be included. Fines and similar deductions should be excluded.

Specified Wage-rates Paid for the Week Ending	Number of Wage Earners Both Time- and Piece-workers Receiving Specified Wage-rates Per Week			
	Adults 18 Years of Age and Over		Young Persons Under 18 Years of Age	
	Males	Females	Males	Females
Under $3 per week . . .	—	—	—	—
$3 to $3.99 per week . . .	—	—	—	—
$4 to $4.99 per week . . .	—	—	—	—
$5 to $5.99 per week . . .	—	—	—	—
$6 to $6.99 per week . . .	—	—	—	—
$7 to $7.99 per week . . .	—	—	—	—
$8 to $8.99 per week . . .	—	—	—	—
$9 to $9.99 per week . . .	—	—	—	—
$10 to $10.99 per week . .	—	—	—	—
$11 to $11.99 per week . .	—	—	—	—
$12 to $12.99 per week . .	—	—	—	—
$13 to $13.99 per week . .	—	—	—	—
$14 to $14.99 per week . .	—	—	—	—
$15 to $15.99 per week . .	—	—	—	—
$16 to $16.99 per week . .	—	—	—	—
$17 to $17.99 per week . .	—	—	—	—
$18 to $18.99 per week . .	—	—	—	—
$19 to $19.99 per week . .	—	—	—	—
$20 to $20.99 per week . .	—	—	—	—
$21 to $21.99 per week . .	—	—	—	—
$22 to $22.99 per week . .	—	—	—	—
$23 to $23.99 per week . .	—	—	—	—
$24 to $24.99 per week . .	—	—	—	—
$25 and over per week . .	—	—	—	—

References

King, W. I. — *The Wealth and Income of the People of the United States.*
Nearing, Scott. — *Income*, Ch. II, pp. 18–52.
Streightoff, F. H. — *The Distribution of Incomes in the United States*, Columbia University Studies, Vol. LII, No. 2.

CHAPTER V

CLASSIFICATION — TABULAR PRESENTATION

I. The Meaning of Tabulation

Progress in understanding or explaining phenomena rests upon the use of scientific method. Similarities and differences must be studied minutely and their causes traced to their foundations. This requires that discrimination and judgment be used according to some clearly defined purpose. In statistical as a part of general method fundamental steps are classification and tabulation.

> "Performed consciously or unconsciously, the act of classification is indispensable to and accompanies every scientific inference. A mind is orderly or slovenly, according as it does or does not habitually and accurately classify the facts with which it comes in contact. The success of an investigation, the worth of a conclusion, are in direct proportion to the fidelity to this principle and the exhaustiveness with which the process is carried out." [1]

Loose thinking and the assignment of cause for effect, or vice versa, result from a denial or a violation of this principle. This truth is involved in all that is suggested in the term "standardization," and applies no less to statistical science than it does to business and economic procedure. It is the principle of orderly arrangement and to violate it is as indefensible when dealing with statistical facts as when for-

[1] Cramer, Frank, *The Method of Darwin: A Study in Scientific Method*, p. 88.

mulating systems of cost accounts, for instance. A cost system which failed to distinguish between overhead and material costs could no more be defended than a statistical summary which grouped together facts of different properties. Combinations must be made on bases that are common. What these are, how inclusive they may be, and what facts are affected by them, can be discovered only through classification.

Classification in statistical methods consists in arranging data into groups according to their common characteristics. Tabulation consists in placing data thus classified into tables — flat surfaces "with breadth not disproportionately small in comparison with length" — which may be read in two dimensions, the items being set opposite the stub (horizontal) and caption (vertical) classifications. Tabulations may be of the first, second, third, or subsequent order, depending upon the amount of detail which they include. Those of the first order contain all of the important details classified according to their most numerous common characteristics. Those of the second, third, and subsequent order contain data in summarized form and are used primarily in text analysis and in specialized studies to focus attention upon some distinctive characteristic which data possess or relationship which they suggest.

Most frequently detailed data are given in the form of appendices or "General Tables," more with the idea of preservation for purposes of record, and as material for intensive and detailed study, than for current or casual use. They constitute the raw material, removed one step from the original entries, to which access is impossible, and are the sources from which special summaries must be made and standards formulated for an appraisement of the grouping and condensing which are made in the summary tables.

Notwithstanding the fact that the distinctions between these forms of tabulations are solely of degree and not of absolute difference, they are important because of the place which each has in the process of analysis and in the presentation of results. The basis of distinction is on the detail included and the amount of grouping and combining used. It is clear that as the grouping and combining process is extended, accuracy and completeness are sacrificed. Just how far this process should be carried and in how detailed a manner the individual characteristics should be portrayed depend upon the character of the original data and the uses which they are to serve. Properly to summarize the detailed facts bearing on complex problems calls not only for statistical sense but also for statistical integrity. To accept all summaries on their face frequently argues either a lack of interest in scientific study or an abundance of ignorance of the delicacy and limitations of the device which has been employed.

Tabulation may also be used to give a synoptical view of numerical facts. In tables of this character no attempt is made to include all data in detailed or in abbreviated form, but only samples chosen at random or according to a fixed purpose. It is in the use of such tabulations that the greatest danger lies and that exercise of the care and scrutiny, discussed above in relation to primary and secondary data, is most imperative. The discrimination necessary to make a representative digest of detailed numerical data presupposes not only breadth of view and intimate acquaintance with all detail, but also the ability to put in short compass the salient facts without unduly emphasizing some factors or attaching too little importance to others. Only in rare cases should conclusive weight be assigned to a digest. It is always wise to acknowledge the limitations of a synoptical

view, and to make frequent references to the detailed tables upon which summaries are based.

II. The Advantages of Tabulation

Of the superiority of classified over unclassified or heterogeneous statistical data in the analysis of economic problems, it seems almost unnecessary to speak. Certain of these, however, may be specified and briefly commented upon.

First: Regularity over Irregularity and the Order of Arrangement.

Order of arrangement in tabulation may be determined by numerical considerations or by time or position conceptions. Great importance is attached to the numerical order in the tables of the publications of the United States Census Bureau where, for manufacturing industries, the amount of capital, the amount of product, value of product, etc., are controlling in the industry and state classifications. In the tabulation of the Wisconsin income tax statistics the average tax per tax payer controls in certain tables, all other data being arranged on the basis of a descending order in this item. Where arrangement is according to the ascending or descending order of a single item, it is unwise to rank the conditions producing the items by the use of consecutive numbers, as 1st, 2d, 3d, etc. The numerical differences are always one, but the frequency differences, to which the numerical scale applies, may be represented either by large or by small amounts. The United States Census, evidently for political reasons, freely employs this device in ranking states and their subdivisions. The illogicalness of the process may conveniently be illustrated by data taken from the Thirteenth Census.

TABLE A

Table Showing the Names of Industries and Numerical Ranking by Value of Product

(United States Census of Manufacturers, 1909)

Industries	Value of Product, 1909				
	Amount	Rank of Industry	Difference		
			Amount	Per Cent	Rank
Leather, tanned, curried, and finished .	$ 327,874,187	18			
Butter, cheese, and condensed milk .	274,557,718	19	$ 53,316,469	19.42	1
Paper and wood pulp	267,656,964	20	6,900,754	2.58	1
Automobiles, including bodies and parts .	249,202,075	21	18,454,889	7.40	1
Smelting and refining lead	167,405,650	30	81,796,425	48.86	9

For value of product, in the instances chosen, a change in rank of 1 is shown to result from an absolute difference, varying from approximately seven to fifty-three and one third millions of dollars, or relatively, by a difference ranging from 2.58 to 19.42 per cent. In one instance, a change in rank of 1 requires five eighths as large an amount as is necessary in another case to occasion a change in rank of 9. In cases where it is desired to rank data according to their ascending or descending order it is far better to reduce them to index[1] or relative numbers, using the beginning, the last, or an average of all, as a base, than to resort to the use of consecutive numbers.

[1] Index numbers are discussed in Chapters IX and X.

Probably the most frequently controlling condition in tabular arrangement is time. When this controls, data, no matter how different in absolute amount or in relative frequency, are chronologically arranged. In many instances this arrangement is unsatisfactory — the time element having no particular significance.

Frequently the controlling factor is contiguity or position. Suppose it is desired to construct a table showing the number of tenant farmers by states in the United States. The table might be arranged according to the frequency of the occurrence of this phenomenon. In this case, undoubtedly, certain of the Southern states would occupy first position. If considerations of relative position or contiguity were made to govern, the states would be listed not according to the frequency of the phenomenon but in the order in which they occur with relation to each other. If South Carolina were listed first, Georgia and North Carolina would follow immediately. Undoubtedly, such an arrangement would be preferable to indiscriminate listing where neither alphabetical, geographical, nor frequency considerations prevail.

Almost invariably, where geographical distribution is a factor in the statistical tables of the United States Census Bureau's publications, the order of arrangement of districts is from east to west, — New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, Pacific. For the number of "Insane in Hospitals on January 1, 1910" this order is numerically roughly descending, for the percentage of population born in other divisions of the United States the order is distinctly the reverse, and for the percentage of population under fifteen years of age it seems to have no significance.[1]

[1] "Insane and Feebleminded," 1910, *United States Bureau of the Census*, Washington, D. C., 1914, p. 18.

The relation between the phenomena described and the controlling fact in presentation — passage roughly from east to west — in these cases is not clear. It would be clear in describing the distribution inland of the European immigrant. Undoubtedly arguments could be advanced for using the reverse order in describing the distribution of the Asiatics in the United States. The point which it is sought to emphasize is that in the determination of the order of tabular arrangement cognizance should be taken, so far as is possible, of the causal relationship or conformity which maintains between the thing and the arrangement of the material used to describe it.

No sacredness inheres in any single order, except it is the alphabetical, but even it has its limitations. The industrial accident rate is not necessarily highest in the "A" states, nor suicides and divorces lowest in the "U" or "W" states. The most emphatic part of a statistical table is its beginning, and normally the order of arrangement should allow the most important detail (measured in terms of frequency) to appear first and permit conformity and causal relationship to be established between fact and representation. If this is done, then the data appear in the table in the relations in which comparisons will be made. More than one consideration, however, may be important. In studying, for instance, mortality rates from tuberculosis, it would be desirable to compare districts in which city congestion is large, yet conditions of climate, of nationality, and of mode of life of those affected would also be important. In such cases the best order will not be one but many. The thing which should *not* control is the absence of any causal or related order, and this frequently occurs where attention is not given to these considerations. Convenience, however, sometimes requires that the alphabetical arrangement control, yet one would not expect the order of the letters to be of real significance

in the distribution of statistical data. If it is given prominence, it should be subsidiary to conditions which are vital.

The following abstracts of tables of different types of statistical data illustrate varying orders. They should be studied to determine what, if any, considerations have controlled the arrangement.

TABLE B

NUMBER OF EMPLOYEES OF RAILROADS IN SERVICE JUNE 30, 1913.[1]

Class	Number
General officers	4,398
Other officers	10,706
Gen. office clerks	84,267
Station agents	37,721
Other station men	167,450
Enginemen	67,026
etc.	etc.

TABLE C

RAILWAY FREIGHT CARS, NUMBER IN SERVICE, 1913.[2]

Class of Car	Number
Box	1,032,585
Flat	147,541
Stock	78,308
Coal	871,339
Tank	8,216
Refrigerator	43,389
etc.	etc.

TABLE D

DEVELOPED WATER POWER RESOURCES, HORSE-POWER, 1900, BY DRAINAGE BASINS.[3]

North Atlantic	Horse-power
St. John River	13,681
St. Croix River	20,500
Penobscot River	70,454
Kennebec River	63,936
Androscoggin River	123,455
Presumscot River	20,569
Saco River	25,332
Merrimac River	161,333
Connecticut River	292,899
Blackstone River	31,435
etc.	etc.

TABLE E

NUMBER OF DEATHS IN THE UNITED STATES BY CAUSES, 1913.[4]

Causes of Death	Number
Typhoid fever	11,323
Malaria	1,565
Smallpox	125
Measles	8,108
Scarlet fever	5,498
Whooping cough	6,332
Diphtheria and croup	11,920
Influenza	7,725
Other epidemic diseases	6,382
Tuberculosis of lungs	80,812
etc.	etc.

[1] *Statistical Abstract of the United States*, 1914, p. 267.
[2] *Ibid.*, p. 266. [3] *Ibid.*, p. 21. [4] *Ibid.*, p. 73.

Second: A Lesser Tax is Placed on the Memory.

Facts which are at all possible of association may much more readily be remembered and compared when logically arranged than when indiscriminately listed. The force of this generalization is keenly felt when one, in order to make a statistical comparison, is required to read page after page of figures laboriously detailed in prosaic form when the same could have been arranged in a table occupying only a fraction of the space and carrying much more emphasis. "In some cases even no attempt is made at tabular presentation. Nine tenths of the expenditure underlying statistical work that sees the light in such form has been wasted, yet some state commissions publish reams of statistics of this nature every year." Illustrating the point, the author of the above says in a note, "Thus the seventh annual report of the Railroad Commission of Oregon, December 15, 1913, contains over eighty pages (pp. 115–237) of closely printed statistical matter presented almost wholly in running text, without tabular arrangement." [1] Rather than being an aid, it is frequently a serious deterrent to have the same facts recited at length without comment immediately following a statistical table. Certainly, it is an expensive and ineffective method of emphasizing that which seems to be of importance.

Third: Visualization of Group Relations is Permitted.

The mere grouping of like with like into a well arranged statistical table permits a rapid survey and a mental picture to be made of data in their related form. This cannot result if they are indiscriminately placed and if they do not constitute when arranged a distinct tabular form.

[1] Parmelee, Julius H., "Public Service Statistics in the United States," *Publications of the American Statistical Association*, June, 1915, pp. 489–505, at pp. 502–503.

Fourth: By Tabular Arrangement Comparisons are Readily Made between Data of Like Character.

The mere placing of closely related items in juxtaposition simplifies comparison and suggests studies which would not otherwise be thought of.

Fifth: By Tabular Arrangement Summation of Items is Facilitated.

Summation may be accomplished without tabular arrangement but at considerable sacrifice of time and effort inasmuch as the items which are to make up the whole are not placed in lines and columns, and one frequently has difficulty in following them. The component parts of totals are not easily recognized without tabular arrangement.

Sixth: By Tabular Arrangement Repetition of Explanatory Phrases, Headings, and Duplicating Items is Reduced to a Minimum.

One frequently sees in public and private reports long drawn out statements of a few simple facts in which the items repeated are numerous and in which considerable expense and time could have been saved had the items been arranged in tabular form. This condition if possible is always to be avoided.

Without attempting to enumerate further specific advantages of tabulation, it may be said that the same advantages in statistical studies, as in other fields of thought, accrue through orderly and systematic arrangement and classification. Classification is a prerequisite for discrimination, and discrimination is essential to scientific study.

III. The Mechanics of Tabulation

Before the actual process of tabulation is begun, it is generally necessary to go through certain preparative steps.

It is almost never possible immediately to transfer data from schedules or other primary records to tabular forms without intermediate steps being involved. This need may be illustrated by considering the tabulation of data relative to occupations and industries. Before tabulation can be begun classifications for both occupations and for industries are necessary. It is impossible to use directly all of the various names under which occupations are listed and to determine offhand the types of industries by the character of the products reported. After occupational and industrial nomenclatures have been reduced to standard form and the classes which are actually to be tabulated determined upon, it is necessary to transcribe the names of the classes directly, or the code numbers which have been assigned to them, on to tabulation cards. Errors of classification and transcription are bound to creep in. To guard against the former, the limits of the classes must clearly be defined, and the conditions governing the entry into them unmistakably outlined. The readiness and the consistency with which individual instances are disposed of depend upon the completeness with which these conditions are realized. To guard against the latter it is frequently necessary to check the accuracy either by testing samples or by "reading back."

The use of tabulating cards makes it possible to list data in their fullest detail, assigning one space to each item and thereby preserving their individuality and making possible any variety of combinations of the items which is deemed necessary. For simple tabulations a plain card ruled into blocks may conveniently be employed. The number of blocks can be adjusted to fit the necessary detail. For more exhaustive tabulations especially prepared cards are available. These are designed for use in mechanical tabulation

machines, the best known of which is the Hollereth. Numerical codes having been outlined to fit the problem, each item may conveniently be listed by number and space on the cards. In using the plain card it is unnecessary in most instances to write in the detail providing a satisfactory code has been employed. A simple mark, such as a cross or a zero for inquiries to which the answer is positive or negative, or which admit of only two classifications, or numbers for more complex groups, may be used to distinguish the facts recorded.

After data have been coded and transcribed on to cards the next process is that of sorting according to the characteristics which it is desired to tabulate. In case a punching machine has been used, the accuracy of the sorting may be checked by holding the cards up to the light and noting whether it passes through the respective holes for the different items. Any obstruction of the light automatically registers an error in sorting. Where mechanical means of sorting or summating are employed the process is done automatically by electrical contact through holes in the cards. Punching machines may be employed to advantage even where electrical machines for sorting or counting are not available. Most generally, however, except in well appointed statistical offices and laboratories, sorting is done by hand.

In comprehensive studies it is best to sort the cards into the more comprehensive groups provided for in the code. Subsequently, each group may again be sorted into as many parts as it is thought desirable to tabulate separately. To illustrate; all cards bearing the code number for native born, for instance, may be sorted into one pile. These may again be sorted into many or few groups, depending upon the detail with which one desires to describe the native born element. The accuracy of the sorting, when done by hand, may be

checked roughly by rapidly turning through the cards and scrutinizing each of them for errors. In order that this may be done conveniently the cards must be relatively small and the edges accurately cut.

After the cards have been sorted the next process is that of counting or summating the frequency of the occurrence of each item. This may be done in connection with the tabular form when direct transcription is made from the schedule or original sheet to the table. When large aggregates must be summated before tabular entry can be made the process is not easy without first listing the facts, and the use of adding machines for this purpose is imperative. It is best for the inexperienced operator to use a listing machine and to retain the listing sheets for future reference. Where detailed tables involving comparisons are to be made, the rough material on the listing paper may subsequently be employed in computing percentages, averages, etc., and also as a basis for new combinations and cross checking.

It is frequently necessary to arrange data into groups and to express the occurrence of each item in a frequency table in the manner described immediately below. In so doing the individual instance *per se* is lost sight of. This need is particularly true respecting data on wages, sales, ages, etc., cases in which it would be difficult, if not impossible, in extensive studies to list each individual instance. The listing or tallying may conveniently be done by arranging the groups into which the individual items are to be placed on the left-hand margin of a sheet of paper and by tallying off opposite each individual group the number of instances occurring. This method has the disadvantage of making impossible any check on the accuracy of the work. An alternative method is that of transcribing the data to be grouped on to small cards and arranging these into groups, thus allowing each

group to be checked by rapidly running through the cards. This method requires all of the data to be copied, thus allowing error to enter from this source. Whichever method is followed the accuracy of the listing should be thoroughly tested before proceeding to the next stage.

IV. The Technique of the Tabulation Form

The technique of the tabulation form suggests such topics as the amount of detail which it is possible and desirable to show in a single table and the structure of tables themselves. Four types of tables may be distinguished on the basis of the amount of detail which they contain. First, is the "single" tabular form. In this type of table one fact only is given importance. The following may be cited as an example:

TABLE F

Table Showing by Years the Number of Real Estate Mortgages in Wisconsin

Year	Number of Real Estate Mortgages in Wisconsin
Total	—
1890	—
1891	—
1892	—
—	—
—	—
—	—

The second type is the "double" tabular form in which two coördinate facts are represented. The following amplification of the single table will serve as an illustration:

TABLE G

TABLE SHOWING BY YEARS THE NUMBER OF REAL ESTATE TAXABLE AND NON-TAXABLE MORTGAGES IN WISCONSIN

YEAR	NUMBER OF REAL ESTATE MORTGAGES IN WISCONSIN		
	Total	Taxable	Non-taxable
Total	—	—	—
1890	—	—	—
1891	—	—	—
1892	—	—	—
—	—	—	—
—	—	—	—
—	—	—	—

The third type is known as the "treble" form, and in this three sets of considerations are brought out. The example below is an amplification of the double type.

TABLE H

TABLE SHOWING BY YEARS THE NUMBER AND AMOUNT OF REAL ESTATE TAXABLE AND NON-TAXABLE MORTGAGES IN WISCONSIN

YEAR	NUMBER AND AMOUNT OF REAL ESTATE MORTGAGES IN WISCONSIN					
	Total		Taxable		Non-taxable	
	Number	Amount	Number	Amount	Number	Amount
Total	—	—	—	—	—	—
1890	—	—	—	—	—	—
1891	—	—	—	—	—	—
1892	—	—	—	—	—	—
—	—	—	—	—	—	—
—	—	—	—	—	—	—

The fourth is known as the "quadruple." In this type four considerations are given expression. The example below is illustrative.

TABLE I

TABLE SHOWING BY YEARS AND BY DISTRICTS OF THE STATE THE NUMBER AND AMOUNT OF TAXABLE AND NON-TAXABLE REAL ESTATE MORTGAGES IN WISCONSIN

YEAR	DISTRICT OF STATE	NUMBER AND AMOUNT OF REAL ESTATE MORTGAGES IN WISCONSIN					
		Total		Taxable		Non-taxable	
		Number	Amount	Number	Amount	Number	Amount
	Total	—	—	—	—	—	—
Total	1st	—	—	—	—	—	—
	2d	—	—	—	—	—	—
	3d	—	—	—	—	—	—
	4th	—	—	—	—	—	—
	—	—	—	—	—	—	—
	—	—	—	—	—	—	—
	Total	—	—	—	—	—	—
1890	1st	—	—	—	—	—	—
	2d	—	—	—	—	—	—
	3d	—	—	—	—	—	—
	4th	—	—	—	—	—	—
	—	—	—	—	—	—	—
	—	—	—	—	—	—	—
	Total	—	—	—	—	—	—
1891	1st	—	—	—	—	—	—
	2d	—	—	—	—	—	—
	3d	—	—	—	—	—	—
	4th	—	—	—	—	—	—
	—	—	—	—	—	—	—
	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—

It will be noticed that the numbers and amounts of taxable and non-taxable mortgages are given both for years and for districts. Chronology is controlling respecting time; and numerical consecutiveness, respecting space. Totals are provided for each year and for all years; for each district and for all districts. The districts are subsidiary to the years in tabular arrangement, the former being repeated under each year and the total for all years, the reason being that it is desired to concentrate attention upon the districts each year, rather than upon the years within each district. Had the latter purpose prevailed, the districts would have been made primary and the years subsidiary in rank. The order of arrangement respecting taxability emphasizes the direct relations between *number* and *amount.* Had the purpose been to emphasize the relation between taxable and non-taxable mortgages, the data involved would have been thrown into juxtaposition under the superior headings "number" and "amount." The order of arrangement should always be that which will best throw into view vital relations and sequences. As noted below, under *Types of Statistical Tables,* the order and arrangement of tabulation forms should make it clear that the significance of data was clearly understood when they were planned.

Of course, more complex tables may be constructed. In fact there are no limits, except those of expense and statistical prudence, to the complexity which tabular forms may assume. It is generally wise, however, to construct several tables to describe complex conditions rather than unduly to burden a single form. The amount of detail that may readily be grasped by the eye is limited, and too great detail often suggests confusion and repels attention. Judgment must be used in this instance as in all aspects of statistical studies. There is no royal road to excellence in table construction,

neither are there hard and fast formulæ to which appeal can be made for guidance in all cases.

Respecting the structure of tables the following general considerations are of importance:

First. The Rulings and Spacings for Major and Minor Headings.

The amount of space assigned to major and minor headings should be in proportion to their respective importances. This may generally be determined by the order in which they appear. Each subsidiary part should be given less prominence than its immediate superior. Likewise, the most subordinate heading should be assigned more space than that given to an individual item in the body of a table. All forms should be set off by double lines at the top and at the bottom. The sides, however, should remain open as they appear on the printed page. By this method distinction is given to the form of the table by the vertical lines in the body being more clearly brought out. Moreover, it is less likely to have a box-like appearance. Major totals should be set off by double lines both horizontally and vertically. Otherwise, as a rule, only single lines should be used. Where a table is complex and is divisible into two or more distinct parts, the separate portions may be set off by double lines. The complexity of form and amount of detail in each case will suggest the wisdom of modifying these general rules.

Second. The Position of Totals.

Until recently, totals in statistical tables were almost invariably placed below the detail which they summate. The Census Bureau at Washington, some years ago, began constructing their tables with totals at the top, and this practice is now quite widely followed. There is much to be said in its favor. The totals so placed are immediately

before the eye and are closely associated with the title. They are almost invariably the items of chief interest, and it is desirable to have them conspicuously placed. With totals occupying this position, totaling is upward and toward the left. The sums of totals in the lines equal the sums of totals in the columns, the check upon the accuracy showing itself in the total at the extreme left and upper corner of the tabular form.

Third. The Suitability to the Page.

Tables, so far as is possible, should be drawn so as to be completed on a single page. In order to do this it is frequently necessary to omit some of the detail or to use a folded insert somewhat larger than the ordinary page. Tabular forms which run from page to page necessitate that headings be duplicated in detail or in such abbreviated form as will allow the order of the columns to be followed. A sufficient abbreviation in some cases is to number the columns so as to correspond with the order appearing on the first page. By the use of inserts this duplication is obviated, and it is usually possible to view a table as an entirety even if long and complicated. This is of distinct advantage and should be striven for in all cases.

Fourth. The Numbering of Columns.

The practice of numbering columns from left to right is not general in tabular forms in publications in the United States. It is characteristic of foreign statistical publications, however, and its use is of distinct advantage in showing the relationship of totals to their component parts and in facilitating references in text treatment. Not infrequently it is necessary, in referring in text analysis to items in detailed tables, to employ awkward descriptive phrases where it would be easy, by citing lines and columns, unmistakably to fix their position. One often hesitates to verify references

because of their uncertainty and the time involved in identifying the items. The costs and inconvenience of numbering both columns and lines are so small, while the value is so material, that it would seem of distinct advantage to adopt both practices in all tables in which the amount of detail is large or the form of the tabular arrangement at all complex. As an alternative to guide or margin numbers — line numbers — some of the United States statistical publications are arranging lines into groups of five. This breaks up the detail and relieves the monotony of an elaborate table, thus making it easier to follow; but it does not solve the difficulties of making detailed references to tables in text analysis and of showing the columns which are summarized into totals. Column numbers are often of real value in helping to interpret relations between columns in a detailed table. These are not always self-evident even to those experienced in statistical study.

V. The Contents of Tables

The contents of tables will always depend upon the purposes for which they are constructed. The first and foremost consideration is that they should bear clearly upon the purposes chosen. Extraneous or unrelated items, which it might be interesting to show, should not be incorporated into a table designed for a distinct purpose. Tables should likewise be easily comprehended both as to purpose and to contents. Any table which calls for considerable study as to the purpose of its construction or the relationships of the items loses much of its value, and sacrifices, in a measure, the purpose for which it is employed; namely, to show clearly and forcibly in classified and tabular form the numerical facts respecting a given phenomenon or condition. The injunction noted above respecting details, the order and

numbering of columns, the position of totals, the suitability of the page, etc., should be remembered in this connection.

Tables should be accurate as to items and totals. Totals are but the functions of the items which compose them, and generally are no more accurate than the items unless errors so compensate each other as to make an accurate picture from inaccurate details. As to whether this condition maintains, one has to satisfy himself by a study of the units employed in the collection of the data, the accuracy of the data themselves, the interpretation assigned them, etc., conditions which are described at length in the sections above referring to Primary and Secondary data. If error is discovered this tends not only to suggest weaknesses in the tabulation method but also to raise a presumption against the accuracy of the details. Totals should be made to cross-check accurately, cognizance being taken of the possibility that compensating errors may appear in both lines and columns and still the cross-check agree. This condition may be guarded against by carefully scrutinizing the items themselves and the position assigned them in the tabular form. A cross-check, however, is not a complete guaranty against inaccuracies within the body of a table.

Bearing upon the question of accuracy is the consideration of the individuality which is submerged in the tabulating process. Abbreviation necessitates that individual items be lost sight of. The amount of grouping allowable depends in all cases upon the character of data and the purposes for which tabulation is used. Grouping is exaggerated in tabulations of the "second," "third," and subsequent orders. These are summaries of details included in those of the "first" order. It is, of course, impossible in most instances to preserve each individual item in all its originality. In all summary tables, however, the sources of data should clearly be

indicated and the manner of their utilization sufficiently detailed so as to guard against incorrect deductions being drawn from them. References should be made to table, column, and line numbers rather than in blanket form.

In most statistical studies there is a certain percentage of data which it is impossible to classify either because of serious omissions, the use of inappropriate, indefinite, or provincial terms, misconstrued inquiries, paucity of data, etc. These residua, if used at all, are generally grouped as "miscellaneous," "not stated," or "unclassified," items. It should always be the aim in tabulation to reduce these classes to minima. Particularly is this true when comparisons are involved and when an undue importance either by including or omitting them might be assigned to unclassified facts. In case they constitute an appreciable part of a whole it is a wise precaution against misunderstanding and a valuable aid in interpretation to add an explanatory note showing in a general way their contents. Normally, such notes do not immediately, if at all, accompany tabular forms. The result of this is generally bad, inasmuch as most people are inclined to overlook the exceptional cases and to accept a table at its face value. As a general rule, statements of the limitations of statistical tables should closely accompany them, be so conspicuously placed that even the uninitiated will see them, and so clearly put that no one but those who purposely ignore them will fail to be governed by their purport. No one is as well prepared to know the limitations of data, at each stage of collection and tabulation, as those who prepare them, and in justice to all their limitations and virtues should clearly be stated. The place for appraisement to appear is where no one can overlook it.

Frequently italics, bold type, percentages, and averages of various kinds are used in detailed tables to emphasize some

outstanding fact or peculiarity. The degree to which this practice is desirable seems to vary inversely with the generality of the table. The functions of summary and "general" tables are not identical. The former are designed largely if not solely for interpretive purposes; the latter to include detail without prejudice of any kind on the part of the compiler. The more nearly these two functions can be kept distinct, the easier it is for the point of view supported in the analytic treatment to be mastered and detailed data to be used by others for the purposes which they may have in mind. Of course, the two cannot always be kept separate. In some cases, particularly in brief studies, the two shade imperceptibly into each other. In fact, in some instances, it may be impossible or unwise to print detailed facts. In those cases both uses may be combined in the same table. But in large and comprehensive surveys differentiation can be made and is desirable. In such studies it is far better to have a complete statement of the limitations of the data, adequate definitions of the units, and reasons for the combinations which are made of them than it is to dispense with these and have the tables bear evidence of finality through nice computations of average and percentage relationships.

It is the purpose of the statistician to make statistical data as comprehensive and full of meaning as they can be made. It is not his purpose, in connection with *detailed* tables, to predigest them. Much time, effort, and money in the writer's judgment are wasted by making a main feature of such tables elaborate net works of percentages establishing varied relationships which the form of the arrangement seems to suggest irrespective, if not in violation, of the logic back of them. To the attentive reader and the investigator not infrequently they are the bases for a legitimate suspicion both as to function and application. To the uninitiated, they

oftentimes seem conclusive and are used in relations foreign to those for which they were intended and disassociated from the detail upon which they are based.

VI. Titles for Statistical Tables

The title of a statistical table should be a brief epitome of the contents. The most important categories should be specifically named but no attempt made to include all of the different facts revealed. This can be done only by a study of the table itself. It is not the purpose of the title to be a complete summary of its contents. It should be short, clearly phrased, well punctuated, and impossible of double meaning. Titles are generally faulty because of omissions, improper phrasings, and inverted order. Normally, the things enumerated in the title should follow the order of the superior and subsidiary headings. For instance, if commanding importance is assigned to wages paid and these are classified according to hourly, daily, and weekly rates, for occupations, and the latter are listed by districts in which found or by the nationalities of those occupied, then this order should be followed essentially in the title. To invert the order is confusing and may be misleading. Illustrations of faulty titles, omissions of column headings, and other details to be guarded against in tabulations might be cited at length but the following will suffice for our immediate purposes. It is not desired to call attention to the statistical errors of any particular publication or organization; therefore references to the sources of the examples are omitted. Each case cited, however, is bona fide. The reader should always be on the lookout for errors and bad form in statistical presentation. In this way he is able to improve his own methods and to benefit by the mistakes of others.

1. *Omissions in Column Headings*

TABLE J

TABLE SHOWING THE CAUSES OF ACCIDENTS RESULTING IN INFECTION

	TOTAL	FATAL	AMPUTATIONS	INFECTED CUTS AND PUNCTURES	INFECTED BRUISES	INFECTED BURNS	INFECTED EYES
Causes of accidents	721	5	4	511	102	53	46
Nails in floor	32	1	—	31	—	—	—
—	—	—	—	—	—	—	—

The above table should have been constructed thus:

CAUSES OF ACCIDENTS	TOTAL	FATAL TOTAL	NON-FATAL					
			TOTAL	AMPUTATIONS	INFECTED CUTS, ETC.	INFECTED BRUISES	INFECTED BURNS	INFECTED EYES
Total . .	721	5	716	4	511	102	53	46

2. *Misplaced and Confusing Headings and Totals*

TABLE K

TABLE SHOWING JOINTER ACCIDENTS REPORTED, BY NATURE OF DISABILITY

GUARDED OR UNGUARDED MACHINES	ALL ACCIDENTS	TOTAL FINGERS CUT OFF	HAND CUT OFF	FINGERS CUT OFF				
				Four fingers	Three fingers	Two fingers	One finger	LACERATIONS OR ABRASIONS
All accidents	77	71	1	4	2	11	27	32
—	—	—	—	—	—	—	—	—

This table should have been arranged thus:

CAUSES OF ACCIDENTS	TOTAL INDIVIDUAL ACCIDENTS	HAND CUT OFF	LACERATIONS	FINGERS CUT OFF				
				Total	Four	Three	Two	One
Total . .	77	1	32	71	4	2	11	27
—	—	—	—	—	—	—	—	—

3. *Faulty Rulings and Misplaced Column Headings*

TABLE L

TABLE SHOWING ACCIDENTS CAUSED BY FALLS OF WORKMEN — BY CAUSE AND DISABILITY

CAUSES OF ACCIDENTS	TOTALS	PER CENT DISTRIBUTIONS	FATAL	LOSS OF FINGERS	INTERNAL INJURIES	FRACTURES	SPRAINS	LACERATIONS	BRUISES	BURNS	INJURED EYES
Total -all Causes . .	1,387	100.0	48	2	30	425	384	110	346	41	1
Falls down	52	3.7	—	—	—	19	15	5	13	—	—
—	—	—	—	—	—	—	—	—	—	—	—

The total columns should have appeared thus:

CAUSES OF ACCIDENTS	TOTAL	
	Number	Per cent Distribution
Total . . .	1,387	100.00

VII. TYPES OF STATISTICAL DATA AND CORRESPONDING TABLES

On the basis of the manner of treatment and the controlling factor in statistical arrangements, tables are of three types. *First*, those which express historical data; *Second*, those which describe a situation or condition in cross-section; and *Third*, those which express variable data of a non-historical character. Each of these types deserves brief consideration.

The controlling factor in tabulations which express historical data is, of course, chronology. Normally, the arrangement is simple and easily comprehended. All of the facts, no matter how diverse in frequency or divergent in type, are controlled by this consideration, thus giving a continuous view from the standpoint of time. This arrangement does not, however, suit all data equally well. Only when a table serves primarily as an instrument of record and when considerations of time are significant should chronology absolutely dominate. In cases where the time element is incidental it should be reduced to a subsidiary position. The degree of prominence to be given to it depends in each case upon the purpose of the table.

The *second* type of tabulation from the standpoint of contents is that in which a situation or condition is described in cross-section. The controlling facts are the relationships which maintain between the respective things described. The following table relating to scales of wages for plumbers in Massachusetts municipalities will serve as an illustration:

TABLE M

TABLE SHOWING UNION SCALES OF WAGES FOR PLUMBERS ON OCTOBER 1, 1913, BY MUNICIPALITIES. (*LABOR BULLETIN* No. 97, MASS. BUREAU OF STATISTICS, p. 39, BOSTON, MASS.)

MUNICIPALITIES	RATES OF WAGES				
	Hour	Day	Week	Overtime (hour)	Sundays and Holidays (hour)
Attleborough . .	0.40\frac{5}{8}$	$3.25	$19.50	0.81\frac{1}{4}$	0.81\frac{1}{4}$
Beverly . . .	.60	4.80	26.40	.90	1.20
Boston	.62$\frac{1}{2}$	5.00	27.50	1.25	1.25
—	—	—	—	—	—

The data refer to a single period of time and reflect the methods of wage payment, among municipalities, and the different rates of wages at the period to which they apply. That is, the table shows not only geographical distribution but also the relationships maintaining between hourly, daily, and weekly wage-rates. For cross-section tabulations of this type commanding importance should be given to those considerations which are most suggestive. Related things should be placed in juxtaposition in order to facilitate comparisons. Before the form is decided upon the relationships which it is desired to emphasize should clearly be determined and the table be prepared to register them. Tabulation is rarely the first step in analysis; frequently it is the last step, the early ones having been taken in deciding upon the form to be used. A large part of the exposition necessary to make plain what it is intended to show can be obviated if a table on its face unmistakably reveals its purpose. There is nearly always a *best* form, and it is the peculiar function of the person using statistics to discover it. After all, tabulation is only a method of summary expression where lines and columns are used to reveal relationships and sequences.

The *third* type of table, from the point of view of its contents, is one which expresses a *variable* fact at a single period of time. In describing a characteristic of a natural phenomenon one is impressed immediately by the regularity which the measurements, in which the characteristic is given, assume. Regularity of distribution around a central tendency approaches the absolute when dealing with numerous samples and with pure chance selection. If one were to compare the lengths of a great number of leaves, chosen at random from a particular tree, he would be impressed by the degree of uniformity and by the regularity of the graduations on either side of those lengths which

might be called normal or typical. The same uniformity of distribution characterizes the stature or weight of men, size of apples, weight of eggs, or of any other natural thing where chance has freely operated in the choice of the samples.

Similar regularity of distribution occurs when one thing is measured many times. The measurements tend to differ because of the limitations of the physical instruments and of judgment in their use, but these tend to be corrected as the number of measurements is increased. That which is typical or characteristic tends to be established, and the exceptions above and below it to become fewer and fewer as the distance from the norm increases.

In the measurements of certain economic phenomena the same tendency toward regularity of distribution as between that which is normal and that which is extreme is noticeable. Wage-rates vary within narrow margins for the same type of labor for a given district, and between districts the differences are not startling. For a given occupation a norm or typical wage tends to be established. Wages above and below this standard may be thought of as exceptional both as to the amounts paid and the number of individuals receiving them. The foot frontage value on a certain residence city street tends to vary only within a narrow margin, the amount of deviation from the extremes being relatively small and the frequencies relatively few. Down-town business blocks tend to be about six to eight stories in height. There are a few blocks higher than twenty stories and a few old-time buildings — misfits — which are but two or three stories high. Most American freight cars have a capacity of from thirty to fifty tons; very few now in use for freight services have a capacity of less than fifteen tons, while few are built with a capacity beyond one hundred tons. The ruling interest rates on real estate mortgages, in Wisconsin in 1904,

were 5 and 6 per cent. Some loans were made at less than 3 per cent; and a few others at more than 10 per cent. The most characteristic rate was 5 per cent. A degree of normality in these examples is noticeable, but it does not maintain generally in the same rigorous fashion in economic as it does in natural phenomena.

TABLE N

FREQUENCY TABLE SHOWING CLASSIFIED WEEKLY WAGES FOR EMPLOYEES IN ALL MANUFACTURING INDUSTRIES IN MASSACHUSETTS, 1912.

(*27th Annual Report*, Statistics of Manufactures of Massachusetts, 1912, p. xxii, Boston, Mass.)

Wage Groups	Number and Per Cent of Employees Receiving Specified Amounts	
	Number	Per cent
Total	681,383	100.0
[1] Under $3 per week	2,266	0.3
[1] $3 but under $4	5,792	0.9
$4 but under $5	16,909	2.5
$5 but under $6	34,070	5.0
$6 but under $7	52,604	7.7
$7 but under $8	63,879	9.4
$8 but under $9	68,787	10.1
$9 but under $10	75,006	11.0
[1] $10 but under $12	103,160	15.1
[1] $12 but under $15	107,677	15.8
[1] $15 but under $20	104,585	15.3
$20 but under $25	32,536	4.8
[1] $25 and over	14,112	2.1

[1] Note the changing widths of the groups and the treatment of the residuum.

TABLE O

Frequency Table Showing the Number of Deaths from All Causes

Registration Area, United States, 1912 (*Mortality Statistics*, 1912, p. 11, Washington, D. C., 1913)

Age of Decedent	Number		
	Total	Male	Female
All ages	838,251	459,112	379,139
[1] Under 1 year	147,455	82,834	64,621
[1] 1 year	29,713	15,748	13,965
[1] 2 years	13,189	6,889	6,300
[1] 3 years	8,240	4,392	3,848
[1] 4 years	6,042	3,178	2,864
[2] Under 5 years	204,639	113,041	91,598
5–9 years	17,274	9,149	8,125
10–14 years	11,436	6,008	5,428
15–19 years	20,343	10,525	9,818
20–24 years	30,997	16,696	14,301
25–29 years	33,762	18,495	15,267
30–34 years	33,743	18,929	14,814
35–39 years	37,916	21,850	16,066
40–44 years	37,885	22,337	15,548
45–49 years	39,624	23,638	15,986
50–54 years	45,496	26,995	18,501
55–59 years	45,732	26,451	19,281
60–64 years	51,097	28,637	22,460
65–69 years	55,492	30,045	25,447
70–74 years	55,650	29,219	26,431
75–79 years	50,772	25,808	24,964
80–84 years	36,678	17,689	18,989
85–89 years	19,559	9,027	10,532
90–94 years	7,082	2,997	4,085
95–99 years	1,493	620	873
[3] 100 years and over	458	169	289
[3] Unknown	1,123	787	336

[1] Note the lower groups. [2] Note the summary of lower groups.
[3] Note the residuum and the "Unknown."

In the statistical treatment of variable phenomena, the frequency table is generally employed. Such a table is constructed by listing singly or in groups and according to their ascending order the units in which a phenomenon or condition is measured, and by arranging opposite them the corresponding frequencies with which they occur. The preceding brief tables will serve as illustrations.

When units of measurements are grouped normally, accuracy of detail is sacrificed, the amount varying directly with the widths of the groups. This, however, depends somewhat on the nature of the material measured. In continuous series the amount depends in large part upon the accuracy of the measurements themselves. By *continuous* series are meant those in which measurements are simply approximations to an absolute value and which differ by small gradations. That is, they are series in which measurements are only approximations, within the limits set up, to an absolute but indeterminate measurement. By *discrete* or broken series, on the other hand, are meant measurements which are determined by the nature of the units in which expressed. In continuous series, measurement is dependent upon the accuracy with which approximations are made. In discrete series, measurements are determined simply by the nature of the units themselves. These considerations may be made clearer if examples of both series are studied. The following example of a discrete series, showing the number of real estate mortgages in Wisconsin in 1904, classified by rates of interest, admirably illustrates the dependence of the frequencies upon units of measurements.

TABLE P

FREQUENCY TABLES SHOWING THE NUMBER OF REAL ESTATE MORTGAGES IN WISCONSIN, 1904, CLASSIFIED BY RATES OF INTEREST

(Constructed from data in *Report of the Wisconsin Tax Commission*, 1907, p. 330)

RATES OF INTEREST	NUMBER OF REAL ESTATE MORTGAGES		
Total	28,961 (a)	28,961 (b)	28,961 (c)
Under 3%	35	35	35
3 and less than $3\frac{1}{2}$%	133	164	133
$3\frac{1}{2}$ and less than 4%	31		1,309
4 and less than $4\frac{1}{2}$%	1,278	1,785	
$4\frac{1}{2}$ and less than 5%	507		10,769
5 and less than $5\frac{1}{2}$%	10,262	10,878	
$5\frac{1}{2}$ and less than 6%	616		10,004
6 and less than $6\frac{1}{2}$%	9,388	9,621	
$6\frac{1}{2}$ and less than 7%	233		4,531
7 and less than $7\frac{1}{2}$%	4,298	4,327	
$7\frac{1}{2}$ and less than 8%	29		1,639
8 and less than $8\frac{1}{2}$%	1,610	1,615	
$8\frac{1}{2}$%	5		60
9%	55	56	
$9\frac{1}{2}$%	1		478
10%	477	477	
12%	2	2	2
16%	1	1	1

A study of the distribution shows that frequencies in groups beginning with the half per cent and extending to

but not including the even per cent are conspicuously less than in those beginning with the even per cent and extending to but not including the half per cent. The relative fewness in the former groups suggests not only a greater concentration on the even than on the half per cent units, but also a greater concentration on the half per cent than on any other fractional units. This is in line with the financial practice of normally calculating interest rates in no smaller fractions than one half per cent units. There is nothing in the nature of the case which requires the units to be continuous and infinitesimally small, and much which requires them to be calculated in larger units and on even numbers. The actual frequencies are determined by the units in which they are expressed and there is no reason for their equal distribution throughout the widths of the groups chosen. As the groups stand in column (*a*), the piling up of the frequencies on the lower side is evident in every case. If they were widened, as in column (*b*), the distribution would still be of the same general character; but the relative degree of concentration on the half per cent and other fractional parts would not be determinable. Column (*b*) is distinctly less suggestive for the separate groups, but distinctly more so for the complete range than column (*a*). By the distribution in column (*c*) — one per cent groups, as $3\frac{1}{2}$ but less than $4\frac{1}{2}$ per cent, etc., — the even per cent in each instance appears in the middle of the group so that the emphasis assigned to it is theoretically distributed over the whole group. This theoretical dispersion does not, however, fit the case; the concentration is still on the even per cents, and any attempt to distribute it evenly over the whole extent is in violation of the facts as revealed in column (*a*). For purposes of subsequent analysis it is often desirable to place the limits of the groups as in column (*c*), but it is always well to remember the actual as distinct from the theoretical distribution.

In fixing the origin and termination of groups in the case of continuous series, it is desirable to assign due weight to the accuracy of the measurements, so as to provide a continuous and uniform distribution of the phenomena through each group. The number of groups chosen is affected by the same considerations, the purpose being to preserve the essential detail of the phenomena as a whole and still to provide for a distribution typical in such cases.

The following table, showing the measurements of the lengths of lobsters, illustrates the point in mind and the difficulties involved in securing a correct distribution, together with the dependence of this upon the accuracy with which the measurements are made.

The measurements are of natural phenomena and there is no reason why they should not be distributed with an approach to regular frequency. In the actual measurements, however, undue prominence is given to measurements falling on the even and half inch units so that the data in the detailed form do not appear to obey any law of regular distribution. A false accuracy is assigned to each measurement and the resulting distribution is very much distorted from that which is characteristic in such cases. Indeed, greater accuracy within the single groups and over the complete distribution may be obtained if the measurements are expressed in wider groups and the resultant frequencies summated to correspond. This has been done in columns (*b*), (*c*), (*d*), and (*e*). The consideration which distinguishes this distribution from that of the mortgage interest rates is the unreal concentration upon even and half inch units in the approximations. In the former case concentration is normal and should be preserved; in the latter case it is fictitious and should be smoothed out by widening the groups. This process in the former case sacrifices accuracy, while in the latter it helps to realize it.

TABLE Q

FREQUENCY TABLE SHOWING DISTRIBUTION OF THE LENGTHS OF LOBSTERS[1]

LENGTHS IN INCHES	(Frequency) (*a*)	½ INCH GROUP (Frequency) (*b*)	¾ INCH GROUP (Frequency) (*c*)	1 INCH GROUP (Frequency) (*d*)	1 INCH GROUP (Frequency) (*e*)
8	6	8			6
8¼	2		11	14	
8½	3	6			
8¾	3				151
9	143	178	181		
9¼	35			474	
9½	241	296			
9¾	55		810		845
10	514	575			
10¼	61			1152	
10½	532	577	638		
10¾	45				1206
11	568	611			
11¼	43		918	920	
11½	307	318			
11¾	11				775
12	414	422	433		
12¼	8			590	
12½	156	168			
12¾	12		489		497
13	321	326			
13¼	5			474	
13½	146	148	153		
13¾	2				579
14	426	426			
14¼			516	516	
14½	90	90			
14¾					370
15	280	281	281		
15¼	1			320	
15½	45	48			
15¾	3		151		152
16	103	104			
16¼	1			117	
16½	13	13	14		
16¾					44
17	30	30			
17¼			33	33	
17½	3	3			
17¾					10
18	7	7	7		
18¼				7	
18½					
18¾			4		4
19	4	4			
20				4	

[1] The measurements in column (*a*) are taken from the *American Statistical Association Publication,* Vol. 7, p. 60. The original data are in a monograph by Dr. Francis H. Herrick on "The American Lobster in the United States," *Fish Commission Bulletin* for 1895.

Groups should invariably be of equal widths. Where this rule is violated false conclusions are likely to be drawn by comparing frequencies.[1] Not only is error likely to result from hasty comparisons of this character but through the employment of unequal sized groups subsequent analysis by approved statistical methods is rendered difficult, if not impossible. The force of this generalization will be more apparent after we have discussed *Dispersion and Skewness.* If for any reason it is desired to change the size of groups, in order to distribute the number of frequencies more in detail, as for instance, in statistics of ages, summaries of the detailed groups should be made and all successive ones be framed in terms of multiples of the narrower ones chosen. The table on the next page showing the distribution of wage-rates of operators in woolen and worsted mills in the United States serves as an illustration of the use of unequal groups and is suggestive of the errors into which one may be led through their use.

Ignoring the widths of the groups and assuming them as equidistant — a very usual thing to do unless one is accustomed to studying such data — it appears that the regular descending order of the frequencies for both male and the total, beginning at the group 10 to 11.99 cents, is abruptly broken at the frequency 2604 for the total, and at 2109 for the males, thus giving a second point of concentration of the wage earners. Of course, the rapid rise of these two instances as well as the retarded decrease in the case of the females is explained by the size of the groups. This table may only rightly be interpreted if full cognizance is taken

[1] See the discussion of this point by Falkner, R. P., in connection with an analysis of "Income Tax Statistics," *Publications of the American Statistical Association*, N. S. No. 110, Vol. XIV, June, 1915, pp. 521–550, at pp. 422, 523, 537. See also the controversy over the meaning of the income tax statistics, published by the Department of Internal Revenue, in *The Annalist*, December 18, 1915, by Carl Snyder, and January 8, 1917, by William P. Malburn, Assistant Secretary of the Treasury.

of the fact that the distribution applies to groups with limits of 2, 5, 6, 10, and 15 cents, as well as to one group which is open at the upper side. If the table had been properly constructed the order of the units — hourly rates of wages — would have been inverted and uniform size groups employed, or groups used which were reducible to multiples of each other. Where it is impossible to use uniform groupings, breaks should be made in the body of the table to call attention to this fact.

TABLE R

FREQUENCY TABLE SHOWING THE NUMBER OF THE OPERATIVES IN WOOLEN AND WORSTED MILLS IN THE UNITED STATES, BY SEX AND BY HOURLY RATES OF WAGES

(*Report of the Tariff Board on Schedule K*, Vol. IV, part 5. House Document No. 342, 62d Congress, 2d session, p. 997)

HOURLY RATES OF WAGES	TOTAL	MALES	FEMALES
Total	30,454	17,343	13,111
75 cents and over	33	33	—
60 to 74.99 cents . . .	60	59	1
45 to 59.99 cents	109	106	3
35 to 44.99 cents	291	287	4
30 to 34.99 cents	486	451	17
25 to 29.99 cents	2,004	1,849	155
20 to 24.99 cents	2,604	2,109	495
18 to 19.99 cents	1,682	1,142	540
16 to 17.99 cents	2,635	2,036	599
14 to 15.99 cents . . .	4,926	3,729	1,197
12 to 13.99 cents	6,007	3,186	2,821
10 to 11.99 cents	6,153	1,453	4,700
8 to 9.99 cents	2,722	757	1,965
6 to 7.99 cents	661	133	528
Less than 6 cents	99	13	86

In writing the limits of groups it is generally well to use no smaller fraction of the whole unit than was employed in the actual process of measurement. For instance, wages expressed in cents would not ordinarily call for a fractional part of a cent being employed to designate the widths of the groups. Likewise, if measurements are made to the nearest half inch the limits of the groups would not normally be indicated by quarter inches. It is generally desirable, in order to guard against confusing the upper limits of a lower group with the lower limits of an upper group, to avoid writing the two in the same form. For instance, the group 30 to 40 may, for convenience, be written 30 to 39.9. In this form it is clear that a frequency of 40 belongs in the group 40 to 49.9. It may not always be so clear in case the limits are expressed in duplicate form.

TABLE S

TABLE SHOWING THE PERCENTAGE RELATION OF THE ASSESSMENT OF PERSONAL PROPERTY TO TOTAL ASSESSMENT

(*Report of the Joint Legislative Committee of the State of New York*, Albany, 1916, p. 260)

RELATION OF PERSONAL PROPERTY ASSESSMENT TO TOTAL ASSESSMENT	NUMBER	WIDTH OF GROUPS IN PER CENTS
Total	53	
Less than one per cent . . .	2	Less than one
From one to three per cent . . .	5	3 [1]
From four to six per cent	5	2 [2]
From six to eight per cent . . .	10	2 [2]
From eight to eleven per cent . .	7	3 [2]
From eleven to thirteen per cent .	12	2 [2]
From thirteen to eighteen per cent .	5	5 [2]
From eighteen to twenty per cent .	3	2 [2]
From twenty to twenty-one per cent	3	2 [1]
Greater than twenty-one per cent .	1	Indeterminate

[1] Upper limit included. [2] Upper limit not included.

The preceding example is illustrative of some of the occasions for confusion resulting from a violation of this principle. In this brief table the second and ninth groups are indefinite in their upper boundaries. According to the way in which they are stated, items of three and twenty-one per cent, respectively, are not to be included, yet it is certain from the succeeding groups that they are included. If they are, the order is an exception to that which characterizes the majority of the groups. As a result, one is left in doubt as to what is intended. Moreover, the groups are so different in size that discredit is thrown upon the whole table.

VIII. Conclusion

A detailed summary of this chapter seems unnecessary. The aim has been to consider only the most important aspects of the subject. The more general phases of classification and their bearing upon scientific method have for the most part been taken for granted.[1] They need no extended consideration in this connection. We have striven only to show the application of classification to statistical facts.

The technique of tabulation has been approached with the problem of the statistician in view, the aim being to call attention to and to warn against certain indefensible practices commonly followed, and, at the same time to formulate as nearly as can be done, rules of general application. Attention is drawn to the characteristic differences in statistical data and to the proper means of bringing them out in tabular form. A logical background is always assumed for the existence of tables, and the reciprocal relation of a point of view and its tabular presentation taken for granted. Tabula-

[1] These are admirably treated in Venn, John, *Empirical Logic*, and in *The Logic of Chance*, as well as in Jevons, W. S., *The Principles of Science.*

tion is always more than a mechanical drawing of lines and inserting numerical symbols. It is analysis by means of facts, numerically symbolized, set out in relation to each other. To its purpose and technique the statistician cannot give too much attention.

References

Bowley, A. L. — *Elements of Statistics*, Ch. IV, pp. 73–107.

—— — *An Elementary Manual of Statistics*, Ch. VI, pp. 50–56.

Durand, E. D. — "Tabulation by Mechanical Means," etc. in *The Transactions of the International Congress on Hygiene and Demography*, 1912, Section Nine, pp. 83–91.

King, W. I. — *Elements of Statistical Method*, Ch. IX, pp. 83–90.

Zizek, Franz. — *Statistical Averages*, Ch. V, pp. 80–91. (This relates to frequency grouping.)

Watkins, G. P. — "Theory of Statistical Tabulation," in *Quarterly Publications American Statistical Association*, December, 1915, pp. 742–757.

CHAPTER VI

DIAGRAMMATIC PRESENTATION

I. Introduction

In the chapter on *Tabulation* our purpose was to emphasize the function of logical classification and arrangement of statistical data. It was learned that primary data must be classified and reduced to order from the heterogeneous form in which they are reported, while secondary data must be rearranged, separated, combined, and worked over to suit the purposes for which they are intended. Respecting both, the essential element in tabulation is classification. The classes into which data fall are arranged logically in the order of importance and placed in lines and columns. Such an arrangement facilitates study, throws related things into juxtaposition, and suggests analysis of facts in their individual and related capacities. Our purpose in this chapter is to contrast tabulation with diagrammatic presentation — the step which logically follows it in statistical studies — and to discuss the value of the various forms of illustration currently used in such studies.

The expression "diagrammatic presentation" is used in a narrower and less inclusive sense than the expression "graphic methods," primarily for the reason that graphs of various types may be used advantageously in connection with averages and other summary expressions. Their functions are so varied and they are susceptible to so many different

kinds of treatment that it seems necessary to distinguish them from mere pictorial illustrations. Generally both are discussed together. We, however, shall distinguish between them and for reasons which will be clearer after we have discussed *Graphic Presentation.*

The purpose of tabulation is to reduce masses of facts to logical order according to the units of measurements in which they are expressed and for the purposes desired. The functions of diagrams are to illustrate these facts according to the order worked out by tabulation. Tabulation is a condition of analysis; diagrams are generally illustrations of conclusions from analysis. The former is necessary in interpretation; the latter are useful in explanation and exposition. Tabulation or classification precedes; the use of diagrams follows. The former generally serves to clarify the meaning of data; the latter frequently to obscure it. Diagrams may never displace tabulation; they may conveniently accompany it if used with discretion. Tabulation alone suggests study and analysis; diagrams alone are more likely to serve as bases for conclusions arrived at without study and to foster a disregard for the details from which diagrams are drawn. Careful analysis of tabulated data is frequently necessary before their full meaning is divulged; a superficial view of diagrams is often gathered upon mere inspection.

Diagrams rarely add *new* meaning to facts which they illustrate. What they do do is to *add to* the meaning by throwing it into relief and by clarifying it. To those who are incapable of interpreting or are unwilling to interpret data in tabulated form they are necessary and at the same time dangerous devices. It is against their superficial and indiscriminate use which we desire to warn the reader.

It is dangerous, as a general rule, to employ analogies in

scientific work, but one may be hazarded in order to show the dependence and secondary character of diagrams in statistical methods. Botanists when classifying plants use established points of distinction to separate them into groups. The common characteristics are noted in detail and become the bases for further study, each sample or group of samples being differentiated from the others by the presence or the absence of chosen criteria. Groups and sub-groups are distinguished and these again are studied in the light of the distinguishing marks chosen. This process is continued until the points of differences are exhausted or until some scheme of organization extending throughout the whole group or groups is discovered. The activities of botanists in classifying plants are analogous to those of statisticians in tabulating data. The common characteristics become the criteria of distinction. The labeling, naming, and mounting of botanical specimens are analogous to illustrating and "mounting," by statistical diagrams, the relations established through tabulation. The former may exist and be independent of the latter in both instances; the latter grows out of and are conditioned by the former in all instances.

What has been said is not meant to detract from the value of diagrams as *aids* in statistical studies. Its purpose has been solely to establish their position and to warn the reader against assigning too great a degree of finality to them or depending upon them to the exclusion of tabulation. Mere illustration is not an end in this case any more than it is in advertising, for instance. Skillfully designed and cleverly drawn pictures may be as necessary to sell an inferior product as highly colored and fanciful diagrams are to attract the interest of the mentally lazy or ignorant, or to drive home a fact to the indifferent reader. If they do this, however, and truthfully present data which they are

intended to illustrate, they serve a real and sometimes a vital purpose. But designs alone are not enough. Diagrammatic illustrations can never replace data themselves, no matter how accurately they tell the truth or how illuminating they are. They are at best statistical aids and should be so viewed by those who use and study them. A well-drawn and cleverly executed diagram is never a guaranty of the value of the statistical facts which it illustrates. The contention which is here made is given substantial support in a recent review of the *Statistical Atlas of the United States.* The reviewer, in questioning the need of such a volume, raised the point of the wisdom of segregating illustrations from tables and from textual analysis. He says:

"Is the policy of segregation a wise one? Presumably these maps and diagrams have had and will continue to have their most effective use in connection with the tables and text with which they were originally published. To place them in a separate volume with the barest textual comment seems unduly to burden the graphic method of presenting facts. Frequently charts and maps greatly strengthen the textual exposition of a subject; they seldom serve as a complete substitute for editorial analysis." [1]

There is a psychology in the use of statistical diagrams which is worthy of brief consideration. The mind is so constituted that it cannot hold at one time a great mass of numerical facts in all their varied relationships. Relations are likely to be obscured in the effort to remember bare figures. Tabulation partly compensates for this limitation. But even when facts are arranged in tabular form, size or magnitude is the only condition which is appreciated. Even this is generally understood in its absolute and not in its relative aspects. The degrees of more or less, with the changes from

[1] Day, Edmund E., Review of "Statistical Atlas of the United States," in *The American Economic Review*, September, 1915, pp. 648–650, at p. 650.

one to the other, expressed for a single time, for a period of time, for a single place, or for an area cannot readily be comprehended when data are in tabular form. The order in which they are arranged may in part compensate for the limitations of tabulation, but cannot entirely overcome them. If, for instance, the order of arrangement is according to magnitude or frequency, as when districts are arranged in the order of the total amount of sales; or where the order is consecutive, as when amounts of loans are listed according to interest rates, an idea of extreme change is readily grasped. The distribution, amount, and frequency of change, however, are appreciated only after they are thrown into relief by some form of diagrammatic illustration. On the other hand, where there is no controlling condition in tabulation, where the order of arrangement is illogical, — or if logical, is not consistently followed, — spatial, time, and frequency considerations, if felt at all, are bound only imperfectly to be comprehended.[1] It is to overcome these imperfections and limitations of tabular arrangement, to introduce devices for showing the proportional relations between facts, and to emphasize the concepts of space and movement, that diagrams of various types are employed.

In tabulation, the power of visualization is only partly realized. True, if tabular forms are properly drawn, data are arranged in lines and columns according to a logical plan. But relations do not stand out. They may be worked out by means of percentages, but at best, in this form, they are abstract. It is not easy to appreciate the degrees of more and less. Comparisons must be made in terms of standards which are themselves abstract. If other

[1] The desirability of having every tabular form determined according to a definite plan and follow a logical order is developed in the preceding chapter, pp. 119–123.

concepts than magnitude are introduced, as, for instance, spatial distribution, the difficulties of making a double comparison out of abstract units are much increased. It is easy to compare absolute differences in interest rates on real estate mortgage loans realized in Illinois with the frequency at which various rates occur, but it is not easy to relate these rates geographically to the several counties of the state without resorting to some form of statistical map. A tabular form in which the counties are arranged alphabetically may be without logical significance. To group the counties by rates may not necessarily be to include contiguous territory. To compare interest rates, amounts of loans, and districts, illustrative diagrams are of great assistance. Even where geographical distribution is not a factor to be displayed, diagrams are helpful in showing relations and sequences.

Probably sufficient has been said to indicate in a general way that diagrammatic illustration adds something to tabulation. Just how this is done and what it is in particular types of illustrations will be made clearer as we discuss the different forms used, the technique of their construction, and the psychological basis upon which each rests.

II. Diagrams for Illustrating Frequency or Magnitude Alone

The diagrams most commonly used to illustrate frequency and magnitude alone are lines or bars, surfaces and volumes, and as a group are known as *pictograms*. Lines or bars are superior to surfaces and volumes, inasmuch as the latter involve relations which are not readily grasped by inspection. For surfaces, the dimensions vary as the square roots of the surfaces; while for volumes, the dimensions vary as the cube roots of the contents. These facts make it difficult correctly

to interpret magnitude, and frequently lead the unexperienced to use illustrations incorrectly proportioned. Instances where this is done are common. In the case of lines or bars the linear dimensions alone are significant, so that relative magnitudes are reflected by proportional lengths.

The following illustrations are introduced merely to make the discussion clear. They are not intended to be exhaustive nor to indicate all of the merits or demerits of the respective methods chosen. The reader, no doubt, has come in contact with other forms and may have devised some which have special merit for the problems with which he is dealing. While there is no one set of standards which can universally be applied, nor one type of illustration that is best under all circumstances, there is much to be said in favor of standardizing more than we have done diagrammatic methods, and certainly of calling attention to devices that may easily be used to deceive. This matter is considered of so much importance that there is now a committee, representing various statistical organizations and engineering societies, studying the problem in all its phases.[1]

Plate 1 is drawn for the purpose of comparing lines, surfaces, and contents when dealing with frequency or magnitude alone. It is clear that absolute differences are much more evident in the lines than in either of the other methods. Only by study is it possible to check up the differences for the surfaces and the solids. Moreover, by casual inspection, relative differences are not exhibited at all by the latter figures.

[1] This committee is known as *Joint Committee on Standards for Graphic Presentation*, and was formed on the request of the *American Society of Mechanical Engineers*. Willard C. Brinton is Chairman. A preliminary report has been published under the title "Preliminary Report Published for the Purpose of Inviting Suggestions for the Benefit of the Committee," in *The Publications of the American Statistical Association*, December, 1915, pp. 790–797.

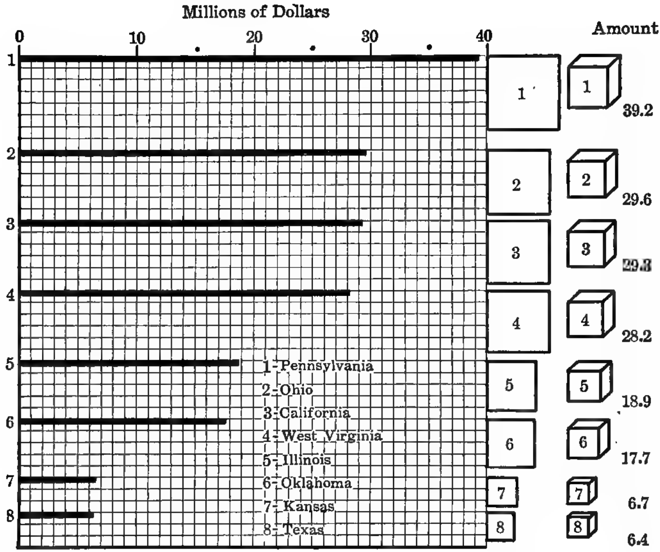


PLATE 1

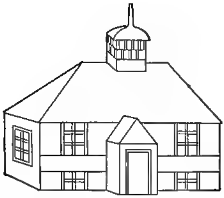
Value of Petroleum and Natural Gas, by States, 1909.
(Illustrations of Lines, Surfaces, and Volumes)

It is only after the square and the cube roots, respectively, have been determined and placed side by side that we get an idea of relation.

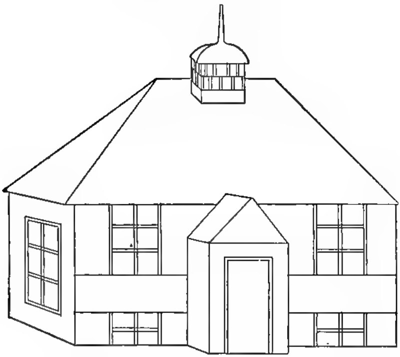
Plates 2 and 3 show solids drawn out of proportion, thus giving erroneous impressions. Such figures are meant to be helpful, but they confuse the reader. In Plate 2, absolute amounts for 1904 and 1914, respectively, stand in the relation of 51.8 to 100. The illustrations show the relation to be 12.5 to 100. In Plate 3, the numerical relation between the amounts is 44.3 to 100; the diagrams show the same to be 6.42 to 100. In both cases, fortunately, the absolute amounts are given, and the errors in the illustrations can be corrected. The latter, considered alone, instead of aiding comparison becloud it.

When it is desired to divide a whole into its component parts, the so-called "pie diagram" is frequently used. It is most popular in showing, for instance, disposition of the parts of a dollar for taxes, wages, interest, profits, etc., and undoubtedly has real merit. (See Plate 4.) Just how superior it is to lines, however, is not clear. Frequently it is necessary to turn the page almost upside down in order to read the legend, and sometimes to insert reference numbers in the sectors because of lack of room for anything more comprehensive. Moreover, for most uses it is more difficult to compare relative sizes in this manner than it is when lines are spread out horizontally before the eye. In addition, the order of presentation is clearer when lines are used. This is evident from the illustration in Plate 5. If diagrams are to be serviceable, they must be easily interpreted. Compare, for instance, the two methods below (Plate 5) of illustrating the petroleum production by states in the United States.

The need for a logical and consistent order of arrangement



1904 $89,282,158



1914 $172,316,862

Per Cent of Increase in 10 Years, 93%

PLATE 2

Public School Property in 1904 and 1914.
(Solids drawn out of Scale)

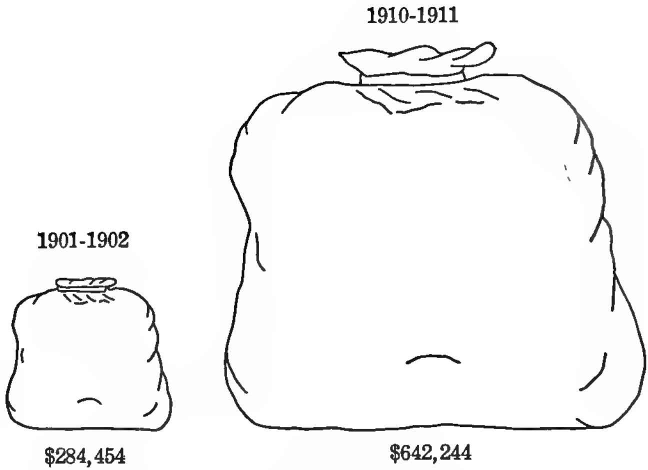


PLATE 3

Payments, Account Bonded Debt and Interest, on County Bonds.
(Solids drawn out of Scale)

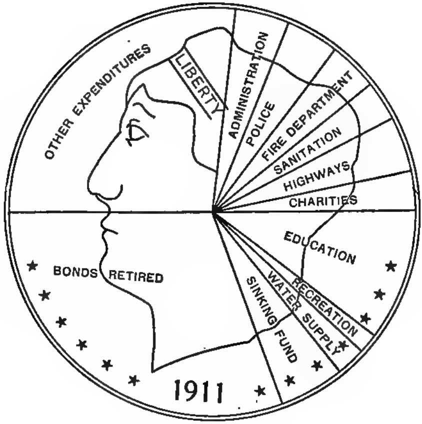


PLATE 4

Our Municipal Expenses, 1911.
(A Pie Diagram)

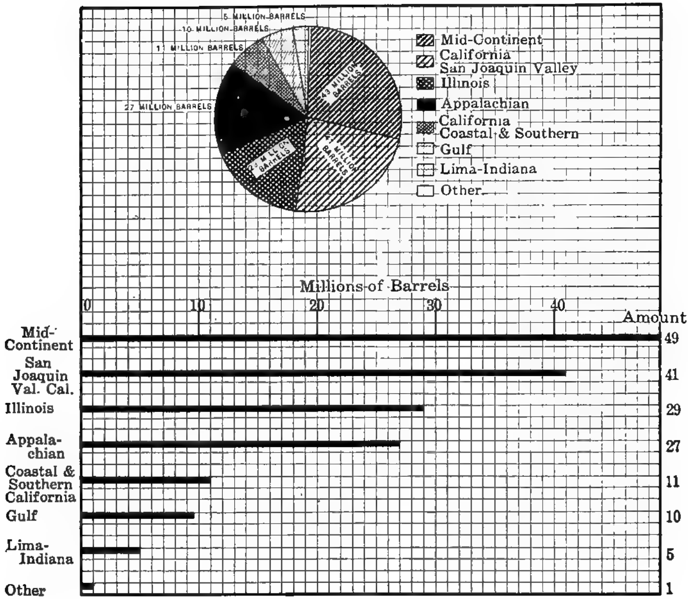


PLATE 5

Production of Petroleum, by Fields, 1909.
(Sectors of Circles and Lines)

in illustration is equally as important as in tabulation. For instance, when dealing with geographical distribution, where contiguity of district is important, this order should be followed. Where time is a factor it should control. The same is true of frequency. As a rule less attention is paid to a logical order of presentation in illustrations than in tabulations for the reason that violations are not generally apparent. False impressions are easily conveyed by the use of an order unnatural to that which the facts normally assume and by omitting all concrete data. Deception, if willed, is not difficult to effect. The apparent is easily confused with the real. It must be remembered that it is the eye and not necessarily the intellect to which appeal is made. And in this very fact lies the chief source of danger in the tendency to think exclusively in terms of illustrations.

Illustrations, whether by bars or lines, surfaces or volumes, ought not to be divorced from the concrete data which they express. The insertion of ordinate and abscissa scales is not enough. Exact magnitudes should be given in illustrations or accompany them in tabular form. When this is done the two supplement and correct each other. The suggestive power of diagrams is not interfered with, and at the same time precaution is taken against the tendency to place reliance in them alone. The failure to include concrete data may not then be used as a partial justification for the drawing of false conclusions. Their presence is a strong deterrent against hasty and unwarranted generalizations and against illustrations being manipulated for illegitimate purposes. The data not only serve as a record of the thing illustrated but also as a test of the accuracy of the illustration.

When lines alone are used their widths are generally without significance. Sufficient space should be allowed so as

to throw into bold relief the devices for distinguishing one set of facts from another. It is, however, necessary when data are classified into unequal-sized groups to use lines of different widths. In such cases it is the surfaces and not the linear dimensions which are important. The widths of lines will vary with the widths of groups but this need cause no confusion if the ordinate scales are properly written, and the surfaces are interpreted in terms of both scales. To depend on abscissa scales alone is inadequate. It is this error which often explains the misinterpretation of data so grouped. An illustration of the erroneous conclusions into which people are led in the use of both diagrams and tabulations by the failure to take into account the changing sizes of groups is given in a recent study of the national income tax.[1] This failure is common and the reader should be constantly on the lookout for it when he is interpreting statistical diagrams.[2]

Frequently, confusion results from including too much in a single diagram, the complexity of detail in whole or in part defeating the functions which it otherwise would have. It is well to keep in mind the general rule that ease of comprehension is a vital consideration and that complex relations can generally more adequately be shown by tabulation. Frequently, however, even for relatively complex relationships, diagrams are of distinct service for the very reason that a number of comparisons can be made simultaneously. For those who are not accustomed to making and interpreting diagrams it is wise to be conservative on the amount of detail crowded into a single figure. There is no general and

[1] See Falkner, Roland P., "Income Tax Statistics," *Publications of the American Economic Association*, June, 1915, pp. 523, 537.

[2] See illustration in *Report No. 4, Industrial Commission of Ohio* on "Industrial Accidents in Ohio, January 1 to June 30, 1914," Columbus, Ohio, 1915, pp. 36–37.

infallible rule respecting this matter, however, since much depends upon the size of illustrations, the skill with which they are drawn, etc.

Plate 6 shows how successfully several facts may be shown on a well-drawn figure. The interesting thing about this figure is that absolute amounts are shown by widths of bars, lengths in all instances being identical and constituting 100 per cent. By cross-hatched surfaces not only are geographical divisions, but color, race, nativity, and parentage shown for the whole population of the United States. The figure admits of being read in two dimensions the same as a table, yet no confusion results. Instead, complex relations are admirably brought out.

When it is necessary to use surfaces and volumes it is best to avoid the placing of areas within areas or contents within contents. If there is a real difficulty in using more than one dimension, it is increased by resorting to this device. It is not clear that such figures should be used except in cases where it is desired to show more than one relation. Even then, by using several illustrations employing lines or bars, the same results may generally be accomplished and with very much less likelihood of misinterpretation and confusion on the part of the reader. In the best statistical publications such figures are seldom used.

Plate 7, showing the adult population in the United States and the number of insane in hospitals, is drawn first in the form of surfaces and second in the form of bars. The first defies comparison. Of course, it is evident that the adult population was greater in 1910 than in 1904, but how much greater is by no means revealed. According to the first method the absolute difference in the number of insane in hospitals at the two periods is barely capable of detection. The illustrations add nothing to the bare facts. So far as

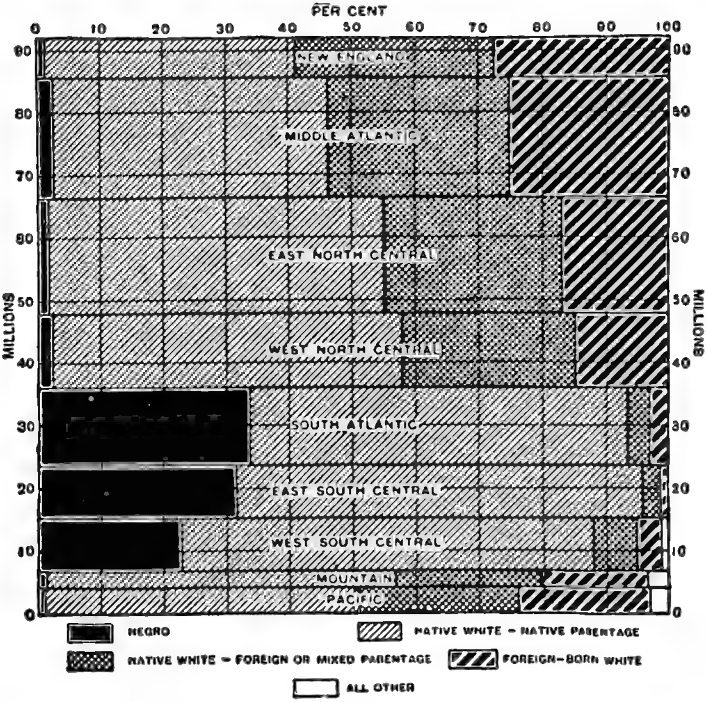
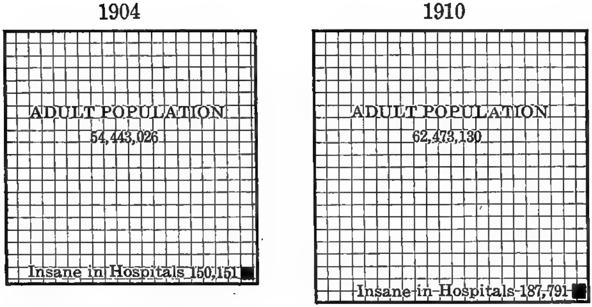


PLATE 6

Color or Race, Nativity, and Parentage, by Divisions of the United States, 1910.



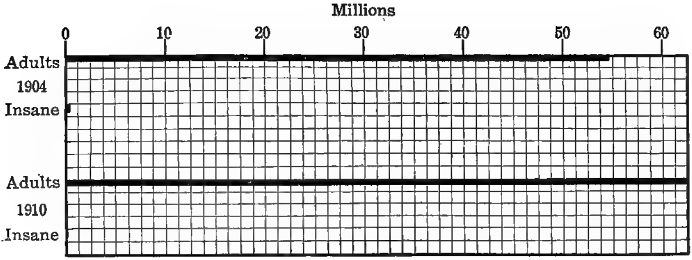


PLATE 7

Proportion of Insane Enumerated January 1, to Adult Population, 1904 and 1910. (Surfaces within Surfaces and Lines)

relations are concerned they are obscured by the manner in which they are shown. Graphically, little aid is given in establishing in either period the relation of the number of insane in hospitals to the total population. An alternative and not very satisfactory method in this case is to use bars.

In summarizing the case for the use of lines and bars in illustrating statistical facts, attention should be called to the appeal which such figures make to the eye and to the ability which they have to make concrete relations and sequences which in tabular form remain abstract. For instance, a hundred per cent becomes significant in a line of a definite length. Likewise, any proportion of this amount is concretely represented by a line somewhat shorter than the one which represents the whole. Undoubtedly, when both the abstract quantity and the pictorial illustrations are employed there results something additional to that which comes from using either alone. It is this something which has its basis in the psychological truth that the intensity with which a thing is perceived varies directly with the number of channels through which it makes its appeal to the intellect.

III. Diagrams for Illustrating Frequency or Magnitude in Relation to Spatial Distribution

1. *The Psychological Bases for the Use of Statistical Maps*

In order to show the relations between magnitude or frequency and geographical distribution various types of statistical maps are employed. They are known as *cartograms* and are in current use in private and public statistical studies. It is our purpose briefly to discuss their psychological bases and to relate them to the principles of statistical methods.

The chief function of statistical maps is to show graphically position in relation to magnitude. For this purpose they are far superior to the tabular form. Data may be spread out geographically and magnitude studied in its relative and absolute aspects. They are likewise superior to simple pictograms, the functions of which are restricted to representing numerical facts according to time and frequency but not according to space. From maps comparisons and contrasts may be made respecting both magnitude and position. The places of absolute and relative concentration and dispersion with the amount and rapidity of change from district to district, near and remote, are thrown into bold relief. Similar comparisons and contrasts are difficult, if not impossible, from tabulations alone. The order of arrangement in tabulation, even if logical and consistent, is fixed and inelastic. Inspection and study may suggest a different order from that chosen but rearrangement is possible only by retabulation.

The order in which data are illustrated on maps, while determined by magnitude or frequency — varying shades of color or density of cross-hatching, etc., indicating varying frequencies — is actually that of contiguity. It is, however, not fixed and inelastic. Comparisons may be made between remote as well as between contiguous districts. Magnitude stands out, being depicted not only alone and in relation to other magnitudes but in relation to position as well. It is the introduction of the spatial concept which is the net advantage of maps over tabular forms and simple pictograms. A new fact is represented — the fact of position — and represented in a different way than it is by tabular arrangement. The order of contiguity may be followed in tabulation, but it lacks the concreteness which the projection upon a map gives it. A new avenue of approach

to the understanding is opened up by statistical maps. It is the approach of visualized position.

Different types of maps reveal the double fact of magnitude and position in different ways depending upon the manner in which they are drawn, and the character of the data which they represent. These are discussed below with their respective merits and demerits.

While maps are superior to tabulations in many ways they are, after all, secondary in character and simply illustrative. Classification and orderly arrangement precede map making. The construction of maps is dependent upon the order, range, and magnitude of data revealed through tabulation and upon the classes into which they fall. In this respect they are not different from pictograms. They do not stand alone. They support and illustrate concrete facts but do not displace them. Hence, they should be accompanied by concrete data, and be interpreted in terms of the units of measurements in which they are expressed. Not infrequently the best that can be done is to show groups into which magnitudes characteristic of districts fall. If groups are wide and magnitudes widely dissimilar, it is impossible even to approximate exact frequency. To guard against misunderstanding, and to validate the form of illustration, maps should be accompanied by concrete facts either directly or in separate tables. Their presence often serves as a positive deterrent to hasty generalizations from appearances, the chief interest being centered on the density of color or cross-hatching and not on the absolute size of the data. In the absence of concrete facts different schemes of illustration may suggest radically different superficial interpretations, since not all types of maps are equally well suited for all purposes. Choice is not a matter to be treated lightly; it is to be determined by the nature and distribu-

tion of the data, the size and character of the groups into which they fall, etc. Maps like simple pictograms are valuable accessories to statistical presentation, but they are not indispensable to statistical analysis.

2. *Types of Statistical Maps*

Classified according to devices for indicating magnitude or frequency, statistical maps are of three general types: those in which frequency is illustrated by different colors or by different shades of the same color; those in which different shades of cross-hatching are used, the frequency or magnitude being indicated by relative densities; and those in which various types of dots indicate frequency.

(1) Colored Maps

The cost of making colored maps is a serious handicap to their general use. Moreover, the superiority of a color scheme over cross-hatching is not always clear. It is sometimes easier to show gradual and minor changes, when groups into which data fall are numerous, by varying the shades of black and white than it is by employing separate colors or different shades of the same color. Changes in color are liable to suggest violent and complete changes in the thing represented, and to accentuate abruptness of change from one condition or district to another. Where different and numerous shades of the same color are used, it is frequently difficult to distinguish between them unless numbers or letters or some other identification marks are used. Color combinations should always be complementary, and shades change in harmony with the facts represented. Lighter colors and shades should represent one extreme; darker colors and shades, the other extreme. On the use of

colored maps, a short extract from "Notes on Map Making and Graphic Representation," by Professor W. Z. Ripley,[1] is of interest.

"It is a cardinal principle in graphic representation that the visual impression should correspond directly to the facts as related to one another. Any scheme of color, therefore, which is not entirely logical, in a visual sense, is worse than misleading when applied to phenomena which are to be represented in a graduated series. A map in which the green, red, yellow, and blue are indiscriminately used to represent different grades of intensity of suicide, for example, is fully as difficult to interpret as the statistical tables which it is intended to elucidate. The only opportunity for representation by means of *unrelated* colors is offered in the case of such phenomena, for example, as the distribution of different nationalities or religions within a country where no relationship in point of fact between the several elements exists. . . .

"If colors are to be used at all, they should either be confined to different intensities of the same color, or else, if the number of shades be too great, two colors, red and blue for example, may be employed, the deepest tints of each standing at the extremes of the series, and each shading down to an almost white color where the two join at the median line."

Numerous and excellent examples of colored maps may be found in the *Statistical Atlas of the United States*, published by the United States Census Bureau, and elsewhere. Those who have occasion to use or interpret such maps should study them in relation to the choice of shades and colors, the varieties of uses to which they are put, the readiness and facility with which they may be interpreted, etc.

(2) Cross-hatched Maps

The second type of maps is that in which some form of cross-hatching is used to indicate magnitude. (See Plate 8.)

[1] *Publications of the American Statistical Association*, Vol. 6, 1898–1899, pp. 313–327, at pp. 314–315.

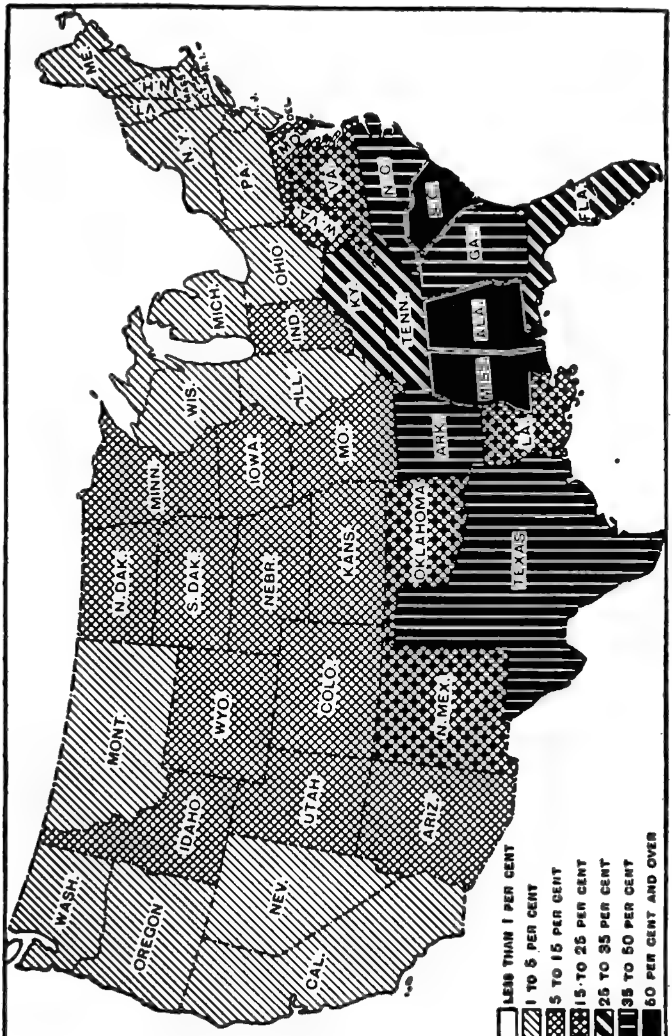


PLATE 8

Proportion of Males 10 to 13 Years of Age Engaged in Gainful Occupations, by States, 1910. (Cross-hatched Map)

Shades may range from white to black, extremes in the range of the thing represented being illustrated by extreme shades, and the condition which is more common, typical, or characteristic by medium shades. The number of shades to be used depends upon the number of groups into which data are divided. As in tabulation, groups should be of uniform size, shades representing equal ranges of units of measurements, rather than equal frequencies with which units occur. The number of times each shade is used in map making, as the frequency with which groups are encountered in tabulation, depends upon the total frequencies represented and the number of shades and size of the groups chosen. As widths of groups in frequency tabulation, so units of shades in cartographic illustration should be uniform. When this rule is followed, choice of shades is of minor consideration. In all cases extreme conditions are shown by extreme shades, that which is typical being represented by medium shades and assuming prominence merely by its preponderance. No confusion need result under these circumstances by arbitrarily changing the shades.

The foregoing discussion applies primarily to the representation of a statistical series. Where unrelated and disassociated facts are illustrated, as, for instance the number of consumers of a given commodity by districts, unrelated shades may be used. In such cases choice is determined largely by the desire clearly to contrast contiguous territories, and at the same time to bring out the detail necessary to the purpose in mind.

Both color and cross-hatching schemes are restricted to data of a "discrete" character. The term "discrete" is used in a somewhat different connection from that in statistical series, yet it is intended to convey similar impressions. In both cases the conditions which fix the limits of

the groups are in a sense predetermined. Where district boundaries are significant either as marking complete changes, the presence or the absence, or the arbitrary limits to the operation of a thing illustrated, as do county or state lines for rates of increase of population, banking facilities, for instance, changes from district to district must appear abrupt and violent. Such maps give the impression that absolute uniformity prevails within districts and that changes occur only between them. For instance, maps illustrating, by districts, per capita sales of merchandise, public revenues and expenses per capita, rates of changes in farm values, rates of increase of crop acreage, the presence or the absence of such a fact as amenability of states to national birth and death registration requirements, the average number of revenue passengers on street and electric railways per inhabitant, etc., must of necessity show conditions as uniform. In such cases relations are dependent upon areas as bases, or upon the presence or absence of a condition which becomes the criterion requiring uniformity of treatment.

If we generalize upon the type of facts which may be shown geographically by systems of cross-hatching and coloring, it is clear that the condition must pertain to the divisions as units and be dependent upon forces which operate within districts. Such maps suggest equal distribution of the phenomenon taking the same color or shade. Breaks appear only at boundaries. There is no attempt to exhibit distribution as a continuous uninterrupted fact. Division lines are predetermined as they tend to be in discrete statistical series. When this condition maintains, this form of illustration is true to the facts. On the other hand, when the fact is subject to gradual change, when it is as necessary to reflect distribution by position within districts as it is between districts, when the forces producing

it are independent of geographical lines, and series are continuous, cartographic representation by abrupt steps at district lines is unreal and gives erroneous impressions. In many respects a more truthful method of illustration of both magnitude and frequency is found in the so-called "dot" maps. This type comprises the third group spoken of above.

(3) Dot Maps

Dot maps may be divided into three classes upon the basis of the kind of dots used. The *first* class is that in which the dots vary in size, each size having a different numerical significance. (See Plate 9.) The scale according to which an illustration is to be drawn having been determined, exact or approximate frequency is indicated in each division of such a map by the number and size of dots. The principle is different from that followed in cross-hatching and coloring. In the case of dots, actual or approximate frequency is indicated within districts; in the cases of both cross-hatching and coloring, only group frequency is illustrated. In the former case, each unit of scale may be represented in each district; in the latter case, only one unit is so represented, the complete scale being shown by the entire map. The determining factor in choice of scale, in the first case, is absolute frequency; in the second case, for matter arranged in series, it is the range of the limits of the measures to which the frequencies apply. Grouping is not provided for in the case of dots and little or no knowledge of geographical distribution is conveyed by exact magnitudes, but only by densities of shades which these magnitudes form. Grouping of frequencies is the cardinal feature of cross-hatched and colored schemes.

As a means of graphically illustrating absolute frequency

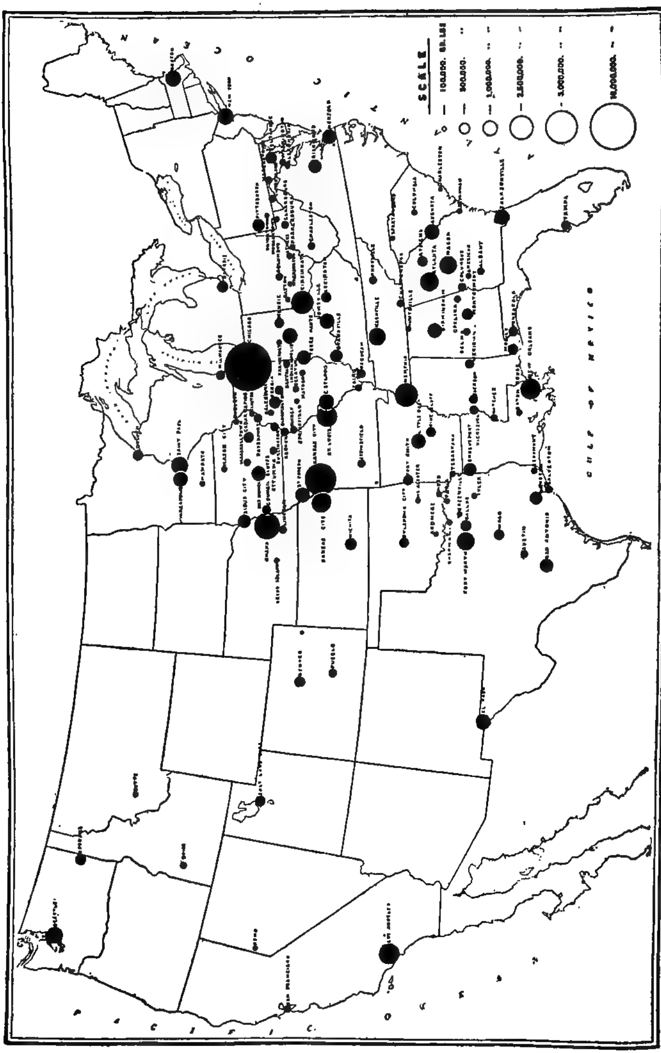


PLATE 9
Primary Markets for Wisconsin Cheese (American) 1911

such maps are failures. It is not evident on inspection, and to determine it involves the double process of counting the dots and relating them to the different values used in the scale. In this respect the method defeats its own end. The process is too tedious and cumbersome. Appeal will be made to tabulation. As a means of roughly indicating geographical distribution they are suggestive, but only in so far as it is done by density of shade. In this particular they add nothing to the ordinary cross-hatched surface. Moreover, they are confusing and may easily be manipulated to give false impressions, inasmuch as surfaces rather than single dimensions are used as bases of comparisons.[1] A circle representing a shipment of cheese of 5,000,000 pounds from Wisconsin to Illinois is not easily compared with one representing a shipment of 1,000,000 pounds into Missouri. Again, they are open to the same criticism as cross-hatching in that they illustrate uniform conditions within and change only between districts. The discussion of this feature respecting cross-hatching applies with equal force to this type of dot maps.

The *second* type of dot maps is similar to the first. Instead of using different sized dots to indicate different steps in the frequency scale, uniform sizes are used, but dots are shaded to indicate different values. (See Plate 10.) Normally, maximum frequency is represented by the solid dot, three quarters, one half, one quarter and other values being shown by variations in the shaded surface. The criticism of the first type respecting varying sizes does not apply in this case, otherwise what is said above in the nature of criticism is of equal significance here. Notwithstanding the fact that they are much in vogue, particularly with the publications of the United States Census Bureau, their

[1] The merits of surfaces and bars are treated above.

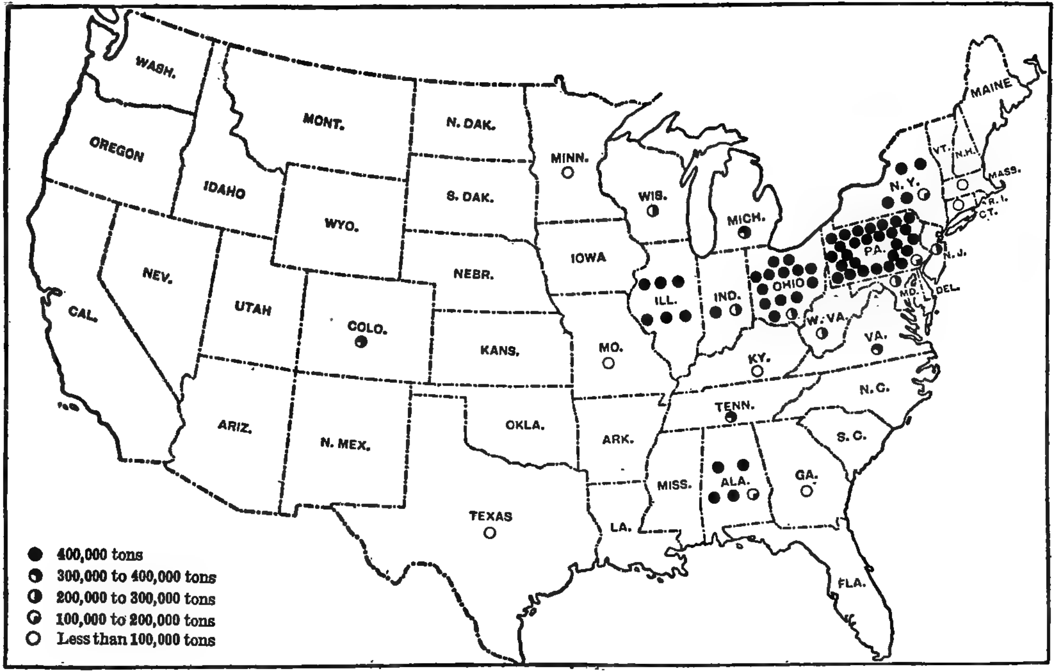


PLATE 10

Pig-Iron Production, by States, 1909. (United States Census, *Statistical Atlas*)

superiority over the old form of cross-hatching for the uses which are common is not proved. In many respects they are at a disadvantage in any such comparison. For other purposes, such as giving a notion of absolute frequency, they add little to the tabular form.

The *third* type of dot maps has decided merits and at the same time certain limitations. The size of the dot is immaterial; the relative frequency with which it occurs is everything. (See Plate 11.) Absolute frequency is secondary, though in theory it may be approximated, as in the other types of dot maps, by considering the number of dots in connection with the value assigned them. Such approximations are generally as unnecessary as they are impossible. Where frequency is great, the number cannot be determined, the individual dots losing their identity in the group. The value assigned to the dot is largely arbitrary, since the purpose of the map is not to record absolute magnitude but to reveal relative abundance and scarcity in relation to position. The densities of the shaded areas are the important facts. Areas of uniform density are not political jurisdictions, as in colored and cross-hatched maps, but actual positions, so far as the sizes of maps will allow these to be shown. This form of illustration gives the impression of gradual changes from scarceness to abundance, from "highs" to "lows," and it seems to smooth out the breaks which would prevail were cross-hatching used. Geographical barriers are ignored in the drawing, but may be inserted for purposes of study and interpretation. It is easy to visualize places and degrees of concentration and "scatteration"; to get a continuous view of distribution. Dot maps of the third type suggest "continuous" rather than "discrete" series.

The technique of diagram and map construction is not here discussed nor even an attempt made to enumerate the

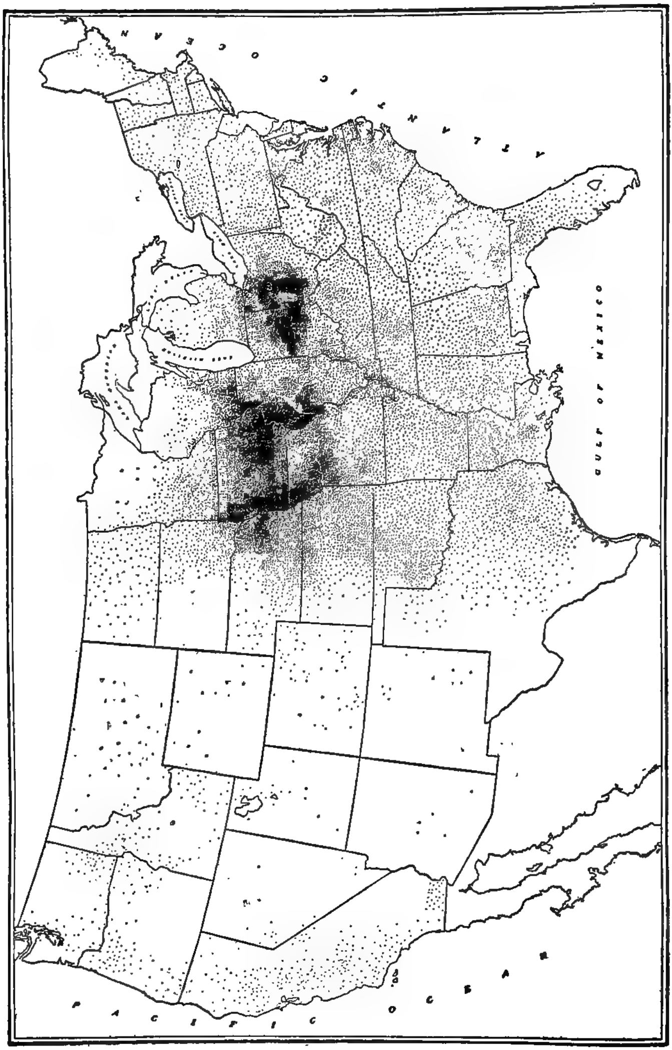


PLATE 11

Number of Swine on Farms and Ranges, April 15, 1910. 1 Dot = 2500

multitude of functions which diagrams serve in the hands of statisticians, publicists, advertisers, manufacturers, financial houses, etc. Numerous examples of well- and ill-drawn illustrations taken from these fields together with a discussion of free-hand and mechanical cross-hatching, the uses of pins in map making, preparation of copy for duplicating whether by photographing or otherwise, etc., are given in *Brinton: Graphic Methods for Presenting Facts.*[1] Our interest is more in describing the functions, discovering and defining the limitations of diagrammatic presentation in statistical studies than in describing the processes of drawing and reproducing diagrams, and in indicating for various businesses the precise functions which they might have in exhibit or other work. Such matters are important but they are treated elsewhere very much more fully than we could hope to do at this time and with all the fullness that they merit.

If the reader understands the psychological bases upon which diagrammatic illustration rests, — if he appreciates the position which it occupies with respect to tabulation and other steps in statistical analysis, and feels the warning, which it has been the purpose of much of the above to sound, against too free a use of or too complete a reliance in pictorial figures, he is in the proper attitude to use the process. Execution may be left to those who have acquired the requisite skill; the determination to use should be in the hands of those who have a correct attitude toward the problem. It is necessary that diagrams should be well drawn and that those who prepare them should have knowledge of the mechanical aids for drawing, duplicating, etc. Such a knowledge constitutes the art; knowledge of the principles underlying the use of diagrams constitutes the

[1] Brinton, Willard C., *Graphic Methods for Presenting Facts*, The Engineering Magazine, New York, 1914.

science, and it is the latter in which we are more vitally interested.

It may be helpful in closing the discussion of the principles and forms of *Diagrammatic Presentation* to outline a few suggestions to be followed in its use.

IV. Suggestions to be Followed in the Use of Statistical Diagrams

1. Choose illustrations which are least liable to be misunderstood, and which most faithfully and correctly interpret the facts.
2. See that fact and representation agree and that all diagrams are provided with concise, clearly stated, and appropriate titles.
3. Avoid figures which must be read according to more than one dimension.
4. Indicate on diagrams the scales of values used, and where necessary to avoid confusion, the dimension or dimensions which are significant in interpretation.
5. Include as a component or as an accompanying part of diagrams the concrete data which they illustrate.
6. In expressing the different parts of a total, use lines or bars or sectors of circles.
7. In statistical maps representing a series, divide the range of frequencies and not the number of districts or divisions into equal parts.
8. In statistical maps representing a series, incorporate as a part of the legend the frequency with which the units of measurements occur, thus indicating the distribution by map and by legend.

References

Bailey, W. B. — *Modern Social Conditions*, pp. 54–56.

Bowley, A. L. — *An Elementary Manual of Statistics*, Ch. V, pp. 35–50.

Brinton, W. C. — *Graphic Methods for Presenting Facts*, Ch. I, pp. 1–20; Ch. II, pp. 20–36; Ch. III, pp. 36–53; Ch. IV, pp. 53–69; Ch. XI, pp. 208–227; Ch. XII, pp. 227–254.

King, W. I. — *Elements of Statistical Method*, pp. 91–97.

——— — *Thirteenth Census of the United States*, 1910. Vol. 5. — *Agriculture, General Report and Analysis. Statistical Atlas of the United States*, 1910 (1914).

CHAPTER VII

GRAPHIC PRESENTATION

I. Introduction

Many of the advantages of diagrammatic apply equally to graphic presentation. The latter deals with graphs or curves of various types which show the distribution of data at a given time or the sequence of data over a period of time. Continuity and relation are emphasized through appeal to the eye, as in the case of diagrams, but more strikingly in that they are uninterrupted.[1] Graphic presentation is beset with many of the limitations which characterize diagrams. The relation to tabulation is secondary; it occupies a subsidiary but frequently a vital position in classification and analysis. Without attempting in any way to repeat the cautions of the last chapter, many of which are applicable to the subject of graphics, we shall assume them and consider the types of graphs, their construction, the conditions under which they may be employed, and the cautions necessary to their use.

There are two types of data which may conveniently be expressed by graphs. *First*, those which at a single instant of time tend to be distributed around a central tendency, and to express the characteristics of a variable fact, and *second*, those which express the occurrence of a homogeneous

[1] Something akin is shown by the frequency type of dot maps. See Plate 11, *supra*, p. 189.

fact or condition over a period of time. In the first instance, the picture is of a fact viewed in cross-section, the measurements being variable; in the second instance, of a fact viewed longitudinally. In the first instance time is of no consequence, degree of change or frequency of occurrence being everything; in the second, time is important, degree of change being expressed in relation to time. A table describing a variable fact and the frequency with which it occurs is called a frequency table, and the curve which describes it a frequency graph. A table which describes the occurrence of a fact over a period of time is known as an historical table and the corresponding curve an historical graph. The graphic presentation of each type of data must be given detailed consideration.

In Chapter V, attention was called to the fact that if a single phenomenon or trait is measured a number of times not one but a number of results is secured. The number of figures a clerk can add in an hour, the length of time it takes to sew a seam of ten inches, the cubic yards of earth which can be removed by a steam shovel in one hour, etc., are variable facts and cannot accurately be measured by a single expression. Completely to describe them the variations must be noted and the number of times which they occur given consideration.[1] On the other hand, phenomena measured not many times, but once, exhibit themselves in a variety of ways and degrees. Some men are tall, others short, cities vary in size, days' work varies in length, wage-rates are frequently widely different, even for the same occupation, salaries are proverbially unequal, freight car-miles per freight train-mile (cars per train), and ton-miles per loaded freight car-mile (tons per loaded car), etc., differ radically for railroads, etc. Such variable phenomena are classified by means

[1] The possibility of reducing a variable fact to a single expression is discussed in Chapter VIII, *infra*.

of frequency tables, *i.e.* tables in which the units of measurements are listed singly or in groups and opposite which are arranged corresponding frequencies.[1] When such a table is graphically illustrated by placing on the horizontal axis — the abscissa — the units or quantities, and along the vertical axis — the ordinate — the corresponding frequencies, we get a surface of frequencies, and when the tops of the ordinates or their middle points are joined together, a distribution curve or graph.

The form and treatment of a frequency graph depend upon the character of the distribution of the variable fact.[2] If measurements are accurately made, if the personal and mechanical elements in their determination are largely removed, and errors tend to be distributed according to chance, large errors will be less common than small ones and the actual measurements tend to arrange themselves around a central or characteristic tendency. This is the case with those distributions which approach the "normal law of error." According to this "law" phenomena are distributed about their averages when the numbers observed are large, and when each phenomenon results from a large number of independent causes none of which is of preponderating importance. Many biological and some economic phenomena, such as the distribution of wages, tend to obey this law. Graphically such series tend to arrange themselves in a bell-shaped figure, the precise shape being dependent upon the degree and place of concentration or scatteration of the frequencies. By no means do all measurements of a variable

[1] See *supra*, Chapter V, pp. 144–156.

[2] On the forms which frequency distributions take see —

Yule, G. U., *An Introduction to the Theory of Statistics*, Chapter VI, pp. 75–105, "The Frequency-Distribution"; Thorndike, E. L., *An Introduction to the Theory of Mental and Social Measurements* (second edition), Chapter III, pp. 28–41, "The Measurement of a Variable Fact."

fact, resulting either from measuring one thing many times or many things once, fall into this regular and "normal" group-distribution. Frequently, there is more than one place of concentration, while at other times no marked central tendency at all appears, both the measurements and their frequencies being widely different.[1] Frequencies may pile up not half-way between the extreme measurements but near one or even both the extremes, the resulting distribution being asymmetrical.[2] If the major concentration is toward the lesser or lower side, the distribution is said to be skewed positively; if toward the larger or

[1] See Plate 23, Chapter XI, *infra*.

[2] The following examples show distributions which are clearly asymmetrical:

Illustration 1

Number of Divorces in the U. S., 1887 to 1906, Classified by Number of Years of Married Life.
(*U. S. Statistical Abstract*, 1913, p. 85.)

No. of Years Married	No. of Divorces
Total	900,584
Under 5 years	255,085
5 to 9 years	282,904
10 to 14 years	162,407
15 to 19 years	91,176
20 to 24 years	54,578
25 to 29 years	29,245
30 to 34 years	15,035
35 to 39 years	6,555
40 to 44 years	2,507
45 to 49 years	805
50 and over	287

Illustration 2

Table Showing Number of Individuals and Corporations Assessed for Income Tax for 12 Wisconsin Counties, classified by amount groups of Assessed Incomes.
(*Rept. Wis. Tax. Commission*, 1912, p. 37.)

Total	11,935
Incomes under $1000	7,890
Incomes $1000 to $1999	1,910
Incomes 2000 to 2999	786
Incomes 3000 to 3999	406
Incomes 4000 to 4999	234[1]
Incomes 5000 to 9999	411[1]
Incomes of 10,000 and over	298

[1] Notice the widths of the groups.

upper side, it is said to be skewed negatively. The measurement of skewness is discussed later.[1] We are now interested in the effect which the form of distribution of the measurements of a variable fact has on its graphic representation.

The distributions of measurements are of two types: *First*, those which form continuous, and *second*, those which form discrete series. A continuous series is one in which measurements are only approximations, within the limits set up, to an absolute value, and which differ among themselves by infinitesimally small gradations. The measurements of

Illustration 3

Table showing Distribution of Percentages of Cost of Collection to Total Collections, Internal Revenue of the U. S., 67 Districts, 1913. (Compiled from the *Report of the Commissioner of Internal Revenue,* 1913, p. 211.)

Percentage Groups	No. of Districts (Frequency)
Total	67
0 to 2	29
2 to 4	24
4 to 6	4
6 to 8	4
8 to 10	4
10 to 12	0
12 to 14	1
14 to 16	1

Illustration 4

Number of Weavers weaving Worsted Goods in the U. S. and Receiving Specified Wage-rates Based upon Actual Weaving Time on Yardage at Regular Piece-rates per Yard, Including Ordinary Stoppage of Loom. (*Report of Tariff Board on Schedule K* — Vol. IV, p. 1007.)

Earnings per Hour	Number
Total	3182
10¢ to 12¢	165
12 to 14	275
14 to 16	375
16 to 18	490
18 to 20	490
20 to 22	438
22 to 24	414
24 to 26	235
26 to 28	150
28 to 30	108
30 to 32	34
32 to 34	4
34 or over	4

[1] See Chapter XI, *infra.*

natural objects belong in this category, since neither size nor weight, for instance, is susceptible to mathematically accurate statement. Age distribution, while generally recorded as a discrete series, is really of the continuous type.

On the other hand, frequencies in discrete series are determined by the character of the units in which the measurements are made. There is nothing in the nature of the case to make them occur at all possible points. Indeed, the nature of the unit determines the points at which the frequencies occur, as for instance, retail prices being expressed in no smaller units than cents; daily wages, in multiples of 25 cents; weight, in no smaller units than pounds; ages, only to the nearest year; express rates in no smaller differences than five cents per pound; passenger fares, in cents per mile, etc. In economic fields the latter series predominates. It is necessary to take cognizance of the types to which series belong when graphically presenting them. Precisely the reason for this being true will be developed in the description of curve plotting. The separate steps to be followed in plotting frequency series of the continuous and of the discrete types will be discussed after the conditions respecting plottings which are common to both, have been described.

II. Graphic Presentation of Frequency Series

1. *Plotting Simple Frequency Series*

Graphically to present a statistical fact two dimensions are used. On the abscissa or horizontal scale are plotted the individual measurements or the groups into which they are put, and on the ordinate scale the frequencies with which each measurement or the combined group of measurements appears. The steps or divisions on both the ordinate and the abscissa axes are represented by equal distances. In order not

unduly to accentuate extreme frequencies, and at the same time to be sure to throw the lesser ones into proper perspective, it is necessary to study the range represented by both measurements and frequencies before deciding upon the scales to employ. Ordinate scales should be made sufficiently small so as to give character to distributions and to allow the frequencies to be determined by reading the curves in terms of the chosen scales. No absolute rule relative to the scales to employ can be formulated.

> "It is only the ratio between the horizontal and the vertical scales that needs to be considered. The figure must be sufficiently small for the whole of it to be visible at once; if the figure is complicated, relating to a long series of years and varying numbers, minute accuracy must be sacrificed to this consideration. Supposing the horizontal scale decided, the vertical scale must be chosen so that the part of the line which shows the greatest rate of increase is well inclined to the vertical, which can be managed by making the scale sufficiently small; and, on the other hand, all important fluctuations must be clearly visible, for which the scale may need to be increased. Any scale which satisfies both of these conditions will fulfill its purpose." [1]

Experience in scale adjustment is the best teacher and a keen sense of form and appearance of the greatest advantage to the student while gaining his experience.

Equal distances on either scale should represent equal facts.[2] The scales should be divided into units which are easily comprehended in terms of the rulings of the paper used. For instance, if paper is ruled in fifths or tenths, the unit of space on the ordinate should be capable of being readily reduced to this basis. Never assign to a space

[1] Bowley, A. L., *Elements of Statistics*, p. 149.

[2] On the necessity of having a horizontal as well as a vertical zero base line, see Clark, Earle, "The Horizontal Zero in Frequency Diagrams," in *Quarterly Publications of the American Statistical Association*, June, 1917, pp. 662–669.

composed of ten small squares such a unit as 3333. Make the space equal to some multiple of ten, as 4000, 5000, 6000, etc. The ordinate scale should be labeled in terms of the arbitrary unit of space adopted and not in terms of the successive frequencies which are to be plotted. Exact frequencies may be inserted opposite the measurements to which they apply if they do not encumber the graph. It is often an excellent plan to insert them horizontally at the top of the sheet on which the curve is drawn.

The abscissa scale should likewise be divided into equal parts. If for any reason successive units are omitted, given in greater detail, or are grouped together into different sized groups, these facts should be made plain by subdividing or widening the unit-area chosen. Under no circumstances should one be left to conjecture as to the precise unit to which frequencies apply. The contention that uniformity in the size of frequency groups is necessary in tabulation has even greater weight when applied to graphic presentation. Assumptions respecting an unbroken continuity are much more likely to be made of graphed than of tabulated distributions.

(1) — Plotting Simple Frequency Distributions Describing Discrete Series

Measurements in discrete series, by custom or otherwise, are expressed in the units in which the thing measured exhibits itself. Illustrations of such series are given above. When they are graphically presented, the units on the abscissa do not represent a tendency the exact measurement of which is impossible to determine because of the limitations of science, or because all possible measurements are likely to occur within the limits set up, but an established fact, subscribing to conditions which can be measured, and according to the customary form in which they are exhibited. The unit on

the abscissa assigned to such a fact, therefore, can almost never be accurately represented by a space. It is almost always a point, and usually the lines connecting the ordinates have no other function than to aid the eye in comparing their respective heights. The lines between the points are significant as to direction but not as to height from the base, since frequencies do not usually occur at these points. If the frequencies with which the express rates, per hundred pounds between various cities, shown in the following table, end in the different integers, were graphically expressed, lines connecting them for each of the numbers, 1, 2, 3, etc., would have no other significance than to give a more definite direction of trend than could be gained from the bare figures.

TABLE A

TABLE SHOWING THE FREQUENCIES WITH WHICH PRESENT AND PROPOSED EXPRESS RATES BETWEEN ST. PAUL AND CITIES NAMED, FOR SHIPMENTS FROM LESS THAN 1 TO 50 LBS. END IN THE INTEGERS

(I. C. C. No. 4198 "In the matter of Express Rates, Practices, Accounts, and Revenues." Opinion 1967)

Integers	Rates Between St. Paul–Minneapolis, Minn., and					
	Sioux City, Ia.		La Crosse, Wis.		Larimore, N.D.	
	Present	Proposed	Present	Proposed	Present	Proposed
1		3		5		4
2		7		5		6
3		4		5		4
4		6		5		6
5	16	4	30	5	19	4
6		6		5		6
7		4		6		4
8		6		5		6
9		4		5		4
0	34	6	20	4	31	6

The same is true of such distributions as the following:

TABLE B

TABLE SHOWING THE NUMBER OF NEW HAMPSHIRE WORKINGMEN IDLE BY WEEKS

(Second Annual Report New Hampshire Bureau of Labor, 1894, pp. 384–385)

WEEKS IDLE	NUMBER REPORTED	WEEKS IDLE	NUMBER REPORTED	WEEKS IDLE	NUMBER REPORTED	WEEKS IDLE	NUMBER REPORTED
1	16	11	5	21	1	31	0
2	60	12	23	22	2	32	0
3	28	13	8	23	0	33	4
4	13	14	6	24	0	34	0
5	*37	15	*21	25	*33	35	*2
6	15	16	6	26	0	36	1
7	21	17	43	27	4	37	0
8	28	18	1	28	0	38	1
9	10	19	0	29	2	39	0
10	*36	20	*15	30	3	40	0

* The starred numbers show the unmistakable tendency to express facts in "round numbers."

TABLE C

TABLE SHOWING THE NUMBER OF FEMALES AND MINORS EMPLOYED IN 24 MERCANTILE ESTABLISHMENTS IN SEPTEMBER, 1913, RECEIVING CLASSIFIED WAGES

("Minimum Wage Legislation in the United States and Foreign Countries" — *Bulletin of the United States Bureau of Labor Statistics* — Whole Number 167, April, 1915, p. 96)

WEEKLY WAGE	NUMBER OF FEMALES AND MINORS RECEIVING SPECIFIED WAGES	WEEKLY WAGE	NUMBER OF FEMALES AND MINORS RECEIVING SPECIFIED WAGES
Total	3,189		
$3.00	20	$14.00	60
3.50	—	14.50	2
4.00	50	15.00	164[1]
4.50	18	15.50	2
5.00	72	16.00	27[1]
5.50	2	16.50	15
6.00	254[1]	17.00	14
6.50	4	17.50	26
7.00	311[1]	18.00	65[1]
7.50	48	18.50	4
8.00	490[1]	19.00	5
8.50	44	19.50	4
9.00	441[1]	20.00	57[1]
9.50	4	—	—
10.00	370[1]	21.00	3
10.50	13	22.00	23
11.00	72[1]	—	—
11.50	8	25.00	37[1]
12.00	355[1]	27.50	7
12.50	16	30.00	9
13.00	22	—	—
13.50	37	35.00	9
		Over 35.00	5

[1] Notice the concentration on even dollar amounts.

In the illustration showing the number of idle men, the unit is arbitrarily taken as the week, and, of course, the corresponding frequencies are at best approximations. The amount of time lost may conceivably be expressed in this manner because of the tendency among employers to lay off men at the close of the week (the pay period) and to take them on at the beginning, yet this practice would hardly account for the wide variation from week to week, and the marked concentration on the fifth week and its multiples. How many people were idle fractional parts of a week, or exactly how much more than a week, is not known, and it is meaningless to attribute significance to the lines which connect the successive ordinates erected at the arbitrary units of measurements.[1]

In Table C, while weekly wages other than those actually named might have existed, it would be an error to suppose that the difference in frequencies between 254 and 4, for $6.00 and $6.50, respectively, were evenly distributed between these two amounts or that there were any persons who received $6.39, for instance. To connect the ordinates representing such amounts is of value only to emphasize the difference and not to establish the distribution between them.

In series in which units of measurements are grouped, while it is customary to represent widths by spaces on the abscissa and to erect ordinates at their middle points, to assume an equal distribution of the instances throughout the

[1] In an analogous case, *The Bureau of Railway Economics*, in plotting the "Monthly Revenues and Expenses per Mile of Line" for the railroads in the United States having operating revenues above $1,000,000, says, "The points on the vertical lines are of significance only in showing the condition for the particular month. The lines connecting the points assist in tracing the change from month to month but do not indicate the trend during the month, nor do they represent cumulative figures for the period." "Revenues and Expenses of Steam Roads in the United States, December, 1915," *Bureau of Railway Economics*, Washington, D.C.

groups unless this is actually the case may lead to serious consequences. A graphic figure should never be accepted as the final criterion of distribution, nor imply a condition which is not realized. For instance, it is known that wage-rates are generally fixed in round numbers, concentration appearing on 5, and its multiples.[1]

To assume even distribution of frequencies within groups of appreciable size for most discrete series is to assume what is either impossible or highly improbable. In many instances, however, such assumptions, though technically incorrect, involve such small margins of error that they are allowable and substantially correct. The validity depends in a large part upon the widths of the groups, on the accuracy of the measurements, and on the regularity and symmetrical character of the distribution.

The following frequency tables emphasize the danger of assuming for discrete series a uniform distribution within groups, such being the result if significance is assigned to straight lines connecting the middle points of ordinates.

[1] Table Showing the Number of Union Bricklayers Receiving Specified Hourly Wage-rates in New York State. (Compiled from the New York Department of Labor Bulletin, Whole No. 65, 1913, pp. 4–6.)

Cents per Hour	Number	Per Cent Distribution
Total	13,362	100.00
50	496	3.71
55	489	3.66
60	1,650	12.35
65	2,391	17.89
70	7,404	55.42
All other	932	6.97

TABLE D

Table Showing the Distribution of Weekly Earnings of Seventy Female Piece-workers Working 50 Hours on Identical Work at Identical Rates, Classified by Groups for One Establishment

(Data are taken from payrolls and valid in all respects)

Division of Groups	Under $3.00	$3.00 to $4.00	$4.00 to $5.00	$5.00 to $6.00	$6.00 to $7.00	$7.00 to $8.00	$8.00 to $9.00		$13.00 to $14.00
First third of the Group . (Exact wages only roughly placed) . .	Notice that this group is three times as wide as the others and that each small compartment corresponds to the	3.09 3.30	4.21 4.28 4.29 4.31 4.32	5.28 5.29	6.03	7.26	8.10		
Second third of the Group (Exact wages only roughly placed) . .	space assigned to each third in the other groups	3.47	4.36 4.40 4.47 4.50 4.50 4.50 4.50 4.50 4.50 4.50 4.50 4.50 4.50 4.50 4.50 4.50 4.50 4.50 4.50 4.50 4.51 4.51 4.51 4.52 4.52 4.52 4.52 4.52 4.52 4.52 4.52 4.55 4.57 4.58 4.66	5.40 5.52 5.55 5.61 5.64		7.43			13.53
Third third of the Group (Exact wages only roughly placed) . .	2.03 2.32 2.68 2.83	3.72 3.87 3.89 3.96	4.76 4.89 4.95	5.70 5.73 5.79		7.92			13.99

TABLE E

TABLE SHOWING THE DISTRIBUTION OF WEEKLY EARNINGS OF SEVENTY FEMALE PIECE-WORKERS BY WAGE GROUPS

WAGE GROUPS	NUMBER OF WAGE EARNERS RECEIVING CLASSIFIED WAGES					
Total . . . . (a)	70 (b)	70 (c)	70 (d)	70 (e)	70 (f)	70 (g)
Under $2.00						
$2.00 to $2.50	2	2	2			2
$2.50 to $3.00	2	5		7	10	
$3.00 to $3.50	3		8			
$3.50 to $4.00	3	11				16
$4.00 to $4.50	8			46		
$4.50 to $5.00	35	38	46			
$5.00 to $5.50	3				53	
$5.50 to $6.00	7	8		11		46
$6.00 to $6.50	1		8			
$6.50 to $7.00		2				
$7.00 to $7.50	2			3	4	
$7.50 to $8.00	1	2	4			4
$8.00 to $8.50	1				1	
$8.50 to $9.00				1		
$13.50 to $14.00	2	2	2	2	2	2

An examination of the weekly earnings in Tables D and E shows how false is the assumption of an equal distribution of frequencies within groups of various sizes. If one-dollar groups above $3.00 are used and these roughly divided into thirds (Table D), not only do the frequencies vary in the same thirds for the several groups, but also in different thirds for the same group. Altering the sizes and limits of groups

does not change matters. As they are widened (Table E), the error in assigning to each possible unit, in which wages might have been expressed, the frequencies indicated by straight lines connecting the ordinates on successive bases becomes all the more apparent. In column (*d*), for instance, which shows the distribution by groups of $1.50, eight persons are shown to receive wages between $5.50 and $7.00, but all of them are in the groups $5.50 to $6.50 and seven eighths in the group $5.50 to $6.00. That is, although the complete group represents three half-dollar groups, one of them is not represented at all in the total frequencies, one by only a single case, and the other by 87 per cent of the total. Widening the groups generally tends to bring regularity out of the complete range of all groups, in case the frequencies follow the "normal" distribution but frequently to sacrifice the accuracy of the details which make it up.[1] If it is dangerous to connect by straight lines ordinates representing frequencies in discrete series, because of implications as to distribution, it is far more dangerous to connect them by smoothed lines on the theory that the distributions follow the ideal or normal type, and that if sufficient samples are taken the irregularities will be smoothed out. If series are discrete, it is this very characteristic which should be retained, and false accuracy is implied in the smoothing process. Only when a smoothed curve gives a more accurate notion of direction and change at successive measures should it be used. It should not be employed as a means of generalizing on the distribution at measures not represented. It is doubtful if the distribution

[1] For examples where successive ordinates in the treatment of wage data are joined together and where the assumption of equal distribution would be dangerous, see "Wages and Regularity of Employment and Standardization of Piece Rates in the Dress and Waist Industry, New York City," *Bulletin of the U. S. Bureau of Labor Statistics*, Whole No. 146, April 28, 1914, *passim*.

of interest rates for real estate mortgages shown in Chapter V[1] would have been materially altered by extending the study over a longer period of time, or by including more instances. Smoothing such curves results in deception. Smoothing may be employed to remove errors in observation but not to disguise the truth. The extent to which it does the latter varies directly for discrete series, with the degree of irregularity characteristic of the thing measured and with the widths of the groups into which the frequencies are forced. (See Plate 12.)

(2) Plotting Simple Frequency Distributions Describing Continuous Series

In plotting continuous series, the contention against joining the ordinates, either by straight or curved lines, loses much if not all of its significance. The fact of measurements being continuous and the units in which they are expressed arbitrary, suggests the propriety of allowing a degree of flexibility for such curves, which for discrete series could not be tolerated. To regard the measurements as accurately and fully descriptive of a continuous series is often as incorrect as to assume all possible measurements for discrete series.

In continuous series, since variations from one extreme measurement to another are regular and gradual, not only should the ordinates be connected, but the direction of the line joining them should be determined by the frequencies at successive and at all measures. Such a curve should be free from sharp angles, the contour being influenced at each point by the relative sizes of adjoining frequencies and by the character of the complete distribution. Let us assume

[1] p. 149.

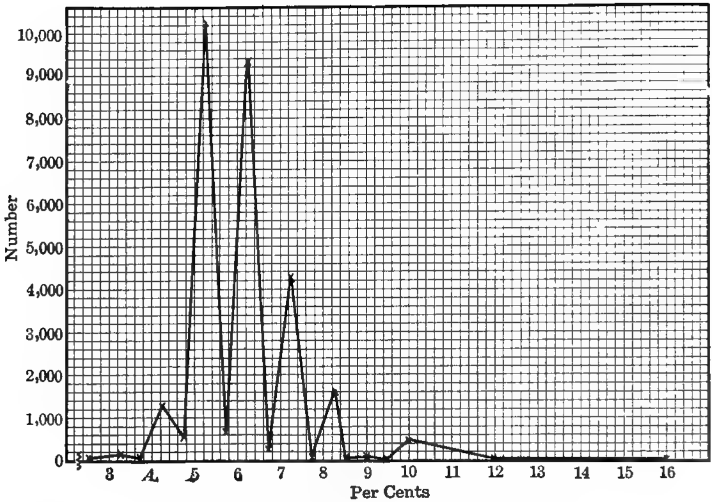


PLATE 12

Number of Real Estate Mortgages in Wisconsin, 1904, by Rates of Interest.
(Frequency Distribution, Discrete Series)

that we were interested in testing the comparative results of planting seed corn from various sized ears and that 327 random sample ears, from seed taken from ears 10 inches long, measured as follows: [1]

TABLE F

TABLE SHOWING THE NUMBER OF EARS OF CORN CLASSIFIED BY LENGTHS

LENGTH OF EARS OF CORN IN INCHES	NUMBER OF EARS AT EACH LENGTH
Total	327
3.0	1
4.5	0
4.0	1
4.5	0
5.0	2
5.5	3
6.0	9
6.5	8
7.0	12
7.5	19
8.0	32
8.5	40
9.0	67
9.5	63
10.0	38
10.5	21
11.0	8
11.5	2
12.0	1

The units of measurements employed have determined the distribution of the frequencies. If they had been more

[1] Davenport, Eugene, and Rietz, Henry L., "Type and Variability in Corn," *Bulletin* 119, *University of Illinois Agricultural Experiment Station*, October, 1907, p. 3.

exact, as for instance, to one tenth of an inch, while the general distribution would have been much the same, the detail would have been distinctly different.[1] To assume that since 40 ears measure 8.5 inches in length and that 67 ears measure 9.0 inches in length, there were no ears with lengths between them — as would correctly be assumed if discrete

[1] "In forming the frequency distribution the measurements are grouped into classes. . . There is no object in taking measurements with extreme accuracy and then grouping them into broad classes. In fact, the nature of the frequency distribution with a given grouping must help to settle the question of grouping, and this in turn the closeness of the measurements. In short, measurements should be so grouped as to show the variability and at the same time to leave the frequency distribution fairly smooth. In the matter of grouping, there are two opposing tendencies — grouping into too few classes to show variability, and grouping into too many classes to give a smooth distribution. In short, the law of distribution is hidden because of too much detail.

"We may lay it down as a general rule that the classes should be only just broad enough to make the distribution fairly smooth, that is, there should be no vacant classes except very near the extremes of the range, and a gradual increase from one extreme up to a maximum and then a gradual decrease to the other extreme, if there is only one maximum in the distribution as is, in general, the case with these populations.

"In respect to grouping into classes the characters treated in this bulletin, we have settled upon one-half inch classes for length of ears, three-tenths inch for circumference, one ounce for weight, and even numbers for rows. This classification or grouping was decided upon after experimenting with classes taken at more frequent intervals.

"There is a further danger of error in grouping besides the narrowness and broadness of classes. For example, at first we measured ears to the nearest tenth inch in length, then suppose we had made quarter inch groupings as follows:

"4, 4.25, 4.50, 4.75, 5.00, 5.25, 5.50, 5.75, 6.00, etc.

"At 5.75 would be grouped all ears which measured 5.7 and 5.8, while at 5.00 would be grouped those which measured 4.9, 5.0, and 5.1. In the long run, this would clearly result in placing more ears at 5.0 than at 5.25, other things being equal. If we should group measurements taken to the nearest tenth inch in 0.5 inch or 0.3 inch classes, no such difficulty arises. Such a grouping as that into quarter-inch groups would not greatly disturb the mean and variability, but would destroy the smoothness of the distribution. Again, if we measure to quarter inches, but group to half inches, some measurements fall on the division lines between classes. Then one half a variate may be recorded in each of the classes between which the variate falls, or if we are dealing with large numbers one can alternately put such a variate into a class above, and below, such a measurement." *Op. cit.*, pp. 27–28.

series were dealt with — is, of course, incorrect. The lines connecting successive ordinates must show no sharp angles, since in the nature of the case, had sufficient samples been taken, ears measuring all lengths between these extremes would have been represented. The same is true of the complete series. While undoubtedly ears essentially 3 and 12 inches in length represent the minimum and maximum, respectively, which would be encountered, the distribution of lengths between these extremes is approximately regular, the degree of irregularity being largely due to the arbitrary units in which the measurements are expressed.

In smoothing graphs of such distributions, effect should be given to the tendency for frequencies, as they approach the maximum, to pile up at the upper side, and as they recede from the maximum to pile up at the lower side, of the groups or measurements in which they are expressed. For instance, in the example used above, between the measurements $7\frac{3}{4}$ inches and $10\frac{1}{4}$ inches, 240 instances are included. The maximum occurs at 9 inches and comprehends 67 instances. At the half-inch measurement below only 40 cases occur, and at the half-inch measurement above 63 instances occur. In the one inch difference between $8\frac{1}{4}$ and $9\frac{1}{4}$ inches, 107 instances are included, 67 of them being in the upper one half. If the measurements were more exact, the unit of difference being smaller, or if the number of samples were increased so as to include all measurements, this piling up would undoubtedly be accentuated. Graphically, this tendency is given expression by rounding the curve to the horizontal as the larger frequencies are approached and rapidly deflecting it to the vertical as the frequencies fall off. Plate 13 shows this fact graphically.

It should be noticed that as the class intervals into which measurements are grouped become smaller, or as the unit-

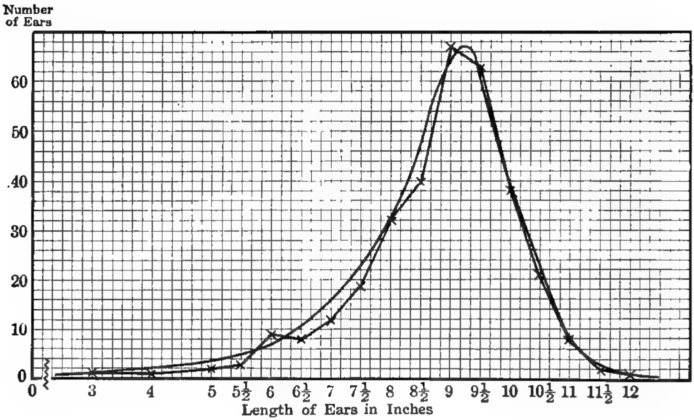


PLATE 13

Smoothed Frequency Distribution of Lengths of Ears of Corn.
(Frequency Distribution, Continuous Series)

accuracy with which the measurements are made becomes greater, and at the same time as the number of observations increases, the lines joining successive ordinates approach smoothed curves. Under different conditions they assume a steplike, halting appearance, unnatural to continuous distributions. In the former case, curves may be smoothed much more readily than in the latter because the exactness and the number of measurements remove the uncertainties under which one works in describing an ideal distribution. A pronounced tendency of distribution, in a continuous series, shown by a fair and adequate number of samples, will tend to be exaggerated if more are taken. On the other hand, if only a few are studied and the resulting curve tends to be very irregular, it is likely that further sampling would result in giving a more characteristic tone to the distribution, making less pronounced both the exceptionally numerous and scarce frequencies. Whether the smoothed curve should exaggerate or give less prominence to extremes depends upon the adequacy of the samples to characterize the distribution of a complete series. No absolute rule can be laid down; the test is the representative character of the samples.[1] Exaggeration or diminution of a tendency should be conditioned by this fact.

2. *Plotting Cumulative Frequency Series*

Up to the present time only simple frequency series have been considered. These are made cumulative when successive frequencies are added together, the result being that the limits of the groups are successively widened. Each

[1] To the rule "that the top of the curve usually overtops the highest point of the frequency polygon, especially when the classes are rather large" (King, W. I., *Elements of Statistical Method*, p. 113), the criticism is pertinent that the determining factor is not so much the size of the groups as it is the representative character of the samples.

frequency class includes all the lower or all the upper ones, depending upon how the cumulating is done. It is immaterial from which extreme measurement the process is begun. If it proceeds from the lesser to the greater, the corresponding frequencies are read "less than," and if from the greater to the lesser, "more than." The following table of prices of oil shows frequencies in both the simple and the cumulated forms, the latter to be read in the "less than" and "more than" manner.[1] It should be noticed that the cumulations are read "less than" when they refer to the upper margins, and "more than" when they refer to the lower margins of groups. For instance, in Table G the number of towns where prices were 10 cents or less was 914; and more than 10 cents, 916. 1830 towns paid 6 cents or more, and 1830, 23.5 cents or less.

TABLE G

Table Showing the Distribution of Towns According to Prices Paid for Oil, Freight Deducted (1830 Quotations), December, 1904, for the United States

(*Report of the Commissioner of Corporations on the Petroleum Industry*, Part II, Aug. 5, 1907, p. 951)

Price, Less Freight (Cents per gallon)	Number of Towns in the United States		
	Simple Frequency	Cumulative Frequency	
		"Less than"	"More than"
Total	1,830	—	—
6.0 to and including 6.5 .	11	11	1,830
6.6 to and including 7.0 .	17	28	1,819
7.1 to and including 7.5 .	27	55	1,802
7.6 to and including 8.0 .	36	91	1,775

[1] The example given is exceptional in that the measurement at the upper margin of each group is *included* in the frequencies. Normally, it is not so included.

TABLE G *Continued*

Price, Less Freight (Cents per gallon)	Number of Towns in the United States		
	Simple Frequency	Cumulative Frequency	
		"Less than"	"More than"
8.1 to and including 8.5 .	123	214	1,739
8.6 to and including 9.0 .	181	395	1,616
9.1 to and including 9.5 .	281	676	1,435
9.6 to and including 10.0	238	914	1,154
10.1 to and including 10.5 .	201	1,115	916
10.6 to and including 11.0 .	162	1,277	715
11.1 to and including 11.5	130	1,407	553
11.6 to and including 12.0 .	85	1,492	423
12.1 to and including 12.5	65	1,557	338
12.6 to and including 13.0 .	49	1,606	275
13.1 to and including 13.5 .	26	1,632	224
13.6 to and including 14.0 .	19	1,651	198
14.1 to and including 14.5 .	43	1,694	179
14.6 to and including 15.0 .	38	1,732	136
15.1 to and including 15.5 .	23	1,755	98
15.6 to and including 16.0	12	1,767	75
16.1 to and including 16.5 .	13	1,780	63
16.6 to and including 17.0	20	1,800	50
17.1 to and including 17.5 .	8	1,808	30
17.6 to and including 18.0 .	7	1,815	22
18.1 to and including 18.5 .	6	1,821	15
18.6 to and including 19.0 .	4	1,825	9
19.1 to and including 19.5 .	1	1,826	5
19.6 to and including 20.0 .			
20.1 to and including 20.5 .			
20.6 to and including 21.0 .			
21.1 to and including 21.5 .			
21.6 to and including 22.0 .			
22.1 to and including 22.5			
22.6 to and including 23.0 .	1	1,827	4
23.1 to and including 23.5 .	3	1,830	3

Cumulative frequencies are helpful in that they furnish continuous summaries of distributions and when reduced to a percentage basis make it easy to determine currently, if the extreme range of distribution is scanned, how one fourth, one half, three fourths, etc., of the frequencies are affected.[1] This is not readily done when one has only the simple frequencies. From the latter, separate pictures of distribution are gleaned, but not a continuous and cumulating photograph. It is as legitimate to cumulate discrete as it is continuous series, so long as the basic distinctions between the two, pointed out above, are kept in mind. The advantages are the same for one as for the other.

When a cumulative frequency series is plotted,[2] the curve may extend from the lower left-hand corner to the upper right, or from the upper left-hand corner to the lower right, depending upon the way in which the cumulating is done. If it is the "less than" form, it follows the first, and if the "more than" form, the second direction. If the former condition maintains, the curve must either be directed upward or to the horizontal; and if the second condition maintains, downward or to the horizontal. In either case, approach to the vertical represents relatively large frequencies and rapid cumulation, and if persistent, a grouping or congregating at this place. That is, the characteristic or modal distribution is revealed by the direction and position of a curve.

In plotting cumulative curves or ogives, as they are often called, the abscissa units, if they represent groups, are indicated as spaces; but if they represent single measurements they are represented as points, the distance between them for discrete series being without significance. For

[1] The use of cumulated frequencies in graphically determining modes, medians, and quartiles is discussed later.

[2] See Chapter VIII, Plates 17 and 18.

simple frequencies in both discrete and continuous series, it is allowable, as has been seen, to plot to the middle points of groups, but the resulting curves must be differently interpreted. Cumulated series are plotted to the upper or the lower side of groups, depending, as has been shown, upon the manner of cumulation. If cumulated frequencies apply to single measures, data for discrete series must be plotted at these points, the lines connecting them giving only the direction or trend. For continuous series, where measurements are so expressed, a straight or smoothed line should be drawn from the middle points of successive cumulations, this being done by assigning on the ordinates vertical spaces proportionate to successive frequencies, and by connecting their middle points. The points to which the lines are drawn are, therefore, typical of the distribution around, and the lines between them typical of the distribution between the measures. Bowley has described such a process, as worked out by Sir Francis Galton, respecting the heights of boys, and it may be helpful briefly to quote him.

"On a horizontal line mark off equal intervals representing units of measurement, say inches.[1] On a vertical scale, mark off equal intervals representing the number of instances, *e.g.*, persons whose heights are measured. Beginning at the lowest, say $51\frac{1}{4}$ inches, on an imaginary vertical line mark as many dots at equal intervals on the vertical scale as there are persons at that height, so that each dot represents one person. From the highest dot thus marked, suppose a horizontal line drawn till it is over the next height division, $51\frac{1}{2}$ inches, and with this new base proceed as before, marking each instance at $51\frac{1}{2}$ inches by a dot vertically above the $51\frac{1}{2}$ inch mark. Next draw a connected line through the middle points of the consecutive vertical rows of dots; if there is an odd number of dots, the middle one is taken as the middle point; if an even number, the middle point is half-way between the middle ones." [2]

[1] The measurements were made to the nearest quarter of an inch.

[2] Bowley, A. L., *Elements of Statistics*, pp. 127–128.

The considerations noted above concerning smoothing simple frequencies of the discrete and continuous types are equally applicable to cumulated series, and do not need further discussion.

Cumulative frequencies and curves are much employed in the business world.[1] They furnish continuous pictures of what has been accomplished in the past and an indication of the direction or trend of future activity. They may be interpreted in terms of both position and slope. When it is desired to make comparisons between different series, it is best to reduce frequencies to a percentage basis, since proportional size of measurement, place of origin and termination of frequencies, and regularity of distribution through the range of measures, can readily be determined by inspection. Whatever their value — and it is frankly admitted to be great — they, like all other graphic representations of statistical facts, rest back upon and are secondary to concrete classified data. There is no desire to belittle their function. Our wish is only to emphasize once more the position which all diagrammatic and graphic representations must hold in the mind of him who uses them in a scientific manner.

III. Graphic Presentation of Historical Series

In graphically presenting historical or time series, the problems encountered are much the same as those found in presenting frequency series. The dimensions are used to represent facts in relation to a constant element — time. There are the problems of choosing appropriate scales, of using a base line, of placing the variable fact on the ordinate, of interpreting the straight or smoothed lines connecting

[1] See Brinton, W. C., *Graphic Methods for Presenting Facts*, especially, Chapters IX and X, pp. 149–163, 164–199.

successive ordinates, of bringing out short- and long-time fluctuations or tendencies, of discovering regularity or irregularity of change, etc. Moreover, there is an approach in many historical, as there is in frequencies series, to what might be called a normal variation. Temperature changes, movements of crops, bank clearings, direction of flow of money, rise and fall of bank reserves, approach regularity with changing seasons or economic disturbances. Crop production and sales of merchandise vary with the amount of rainfall; exports and imports, immigration and emigration, building construction and demand for products of mill and factory, increase and decrease with periods of boom and depression.

It is the problems of expressing these phenomena graphically, of bringing out the short- and long-time tendencies, both absolutely and relatively, with which we are now concerned. Tables describing the occurrence of a variable fact over a period of time are known as historical tables, and the corresponding curves, historical graphs or historigrams. It is the latter with which we are now dealing.

1. *Plotting Simple Historical Series*

The chief problems in the technique of plotting historical graphs relate to the showing of absolute or ratio differences, the necessity of having a base line, the types of lines to use to connect successive ordinates, the purposes and methods of smoothing historigrams, simple as contrasted with cumulative graphs, etc. Each of these is discussed, some briefly and others more fully.

(1) Choice and Adjustments of Scales

In choosing scales for historigrams, in order to show absolute differences, it is necessary to study the extreme range

of variations and to adopt that unit of measurement which neither overaccentuates nor minimizes extreme fluctuations. What the scale will be in a given case will depend, among other things, upon the size of the page, the ability of the eye to view the illustration as a whole, its subsequent treatment, etc. If a single curve is to be smoothed by having another superimposed upon it, the scale should be sufficiently large so as to admit the peculiarities of both to be seen. There is no rule here, as there was none respecting frequencies, which will suffice for all occasions. The most appropriate scale may have to be determined by trial at first, but as experience is an excellent teacher, the trial and error method will not long have to be depended upon.

It is always desirable to plot the variable factor on the ordinate axis and to begin the measurements from a zero base line. If this is not practicable, attention should be called to the fact by drawing a wavy line (∼∼∼) parallel to and slightly above the axis of abscissa. As in the case of frequency series, the ordinates, rather than the range of variates, should be divided into equal parts, and values, which are multiples of the number of spaces into which the paper is ruled, be assigned to each. Equal periods on the abscissa, likewise, should be indicated by equal spaces.

In case two or more curves are to be shown on a single sheet, and they are to be compared in any way, it is frequently necessary to adjust the scales for the different quantities or values indicated. When one curve is a component part of another, the ordinate unit may remain the same, the absolute or relative difference being evident from their positions on the ordinate. If they are widely different, the two may be thrown closely together by adhering to the same ordinate scales but by indicating a break by means of a wavy line drawn between them parallel to the base. If they are related

and expressed in the same unit, the absolute difference being large, the scale may be reduced to a comparable basis by scale conversion.

One method of scale conversion may be illustrated as follows: It is desired to compare graphically the capital and clearings of the New York Clearing House banks. The capital is expressed in millions and the clearings in billions. The absolute difference makes it difficult, if not impossible, to use a common ordinate scale since the curves would be too far apart. They, however, may be brought closely together by equating the scales on the basis of their respective averages. The average capital, for the period 1902–1915, is 140 millions, and the average clearings for the same period 89 billions. These stand in the ratio of 1 to 640. If scales are adjusted, as in Plate 14, so that the ordinates for the two factors stand in this relation throughout the whole period, and amounts are plotted, the curves are thrown closely together and their general direction may be studied. Doing this amounts to plotting the differences of the items from their respective averages, and requires that each curve be interpreted in terms of the unit of equivalence. Currently this is rather difficult to do.

A less common method of bringing together related data, widely different in absolute amount, is by equating on a scale the averages of the deviations (differences) of items from their respective averages and by plotting these deviations and not the original data. This method is more frequently employed when it is desired to give a mathematical expression to these differences than to compare the absolute quantities of the series in question.

More common methods of scale adjustment are those of converting individual variables into percentages of a total, and of expressing them in the form of index or relative

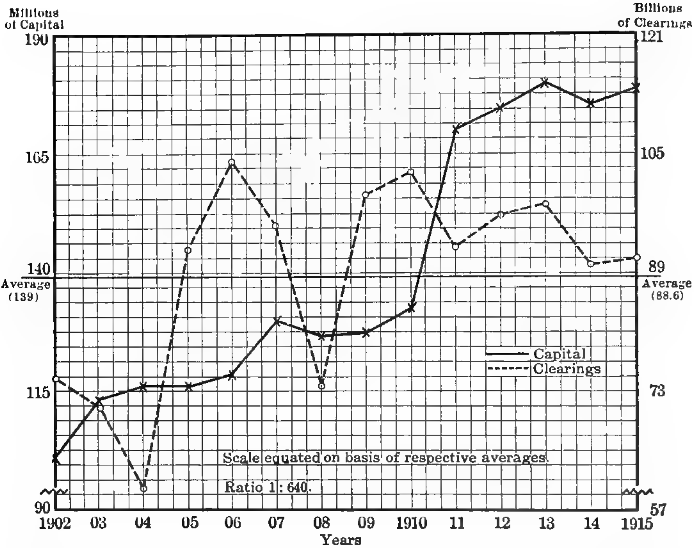


PLATE 14

Capital and Clearings of New York Clearing House Banks, 1902–1915.
(Method of Scale Conversion)

numbers.[1] The first is very common, particularly when absolute differences are large and it is desired to bring curves closely together. It must be remembered, however, that relative and not absolute differences are shown, that they are to be interpreted with respect to each other in the same series, and not in the different series, and that the curves do not necessarily begin nor end at the same point on the ordinate. On the other hand, if the index number or relative method is used (see Plate 15), while variables are expressed as percentages, the base upon which they are computed is not a total but the first, the last, or an average of the different variables. Of these alternatives the last under certain conditions[2] is undoubtedly superior. Of the other two, the last variable (thought of chronologically) is the better base since, other things being equal, one has greater interest in a near than in a remote period, and since the difference in per cent, for successive variables, is more readily calculated from a recent 100 per cent. When this is done, rates of increase and decrease are comparable, equal percentage increments and decrements being represented by equal changes in the ordinate. This method has the disadvantage of beginning or ending the curves at the same points on the ordinate if the first or the last variable is taken as the base, which is sometimes confusing, but the advantage of placing the graphs in close proximity and of registering the general direction or trend from a common start or close. There is the further disadvantage that the values are relative and not absolute, but this can partly be overcome by including the original as well as the percentage data on the graphic figure. Care should be used in reducing absolute amounts to such a basis,

[1] Index numbers are discussed in Chapters IX and X.
[2] These are discussed for index numbers of prices in Chapter IX.

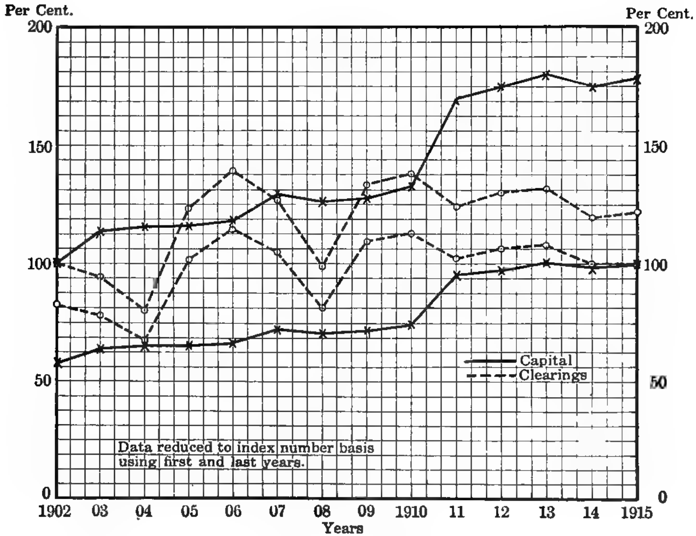


PLATE 15

Capital and Clearings of New York Clearing House Banks, 1902–1915.
(Method of Scale Conversion)

since for many uses the absolute and not the relative changes are significant. The opposite, of course, is likewise true.[1] Percentage or relative figures are always dangerous when the bases upon which they are computed are widely dissimilar, as, for instance, when price increases are compared. A price increase of 50 per cent for a low-priced commodity infrequently used would have little in common, except the nominal increase, with a price increase of the same amount for a high-priced commodity entering into daily consumption on a large scale. One might look with perfect equanimity upon an increase of 100 per cent in the price of lawn seed, but seriously object to the same percentage increase in the price of beefsteak.[2]

(2) The Treatment of Lines Connecting Successive Ordinates

The ordinate scale having been decided upon so as properly to bring out the absolute or ratio differences in a series, the next problem is the determination of the abscissa units and the treatment of the ordinates raised upon them. In historical series the periods of time generally represent accumulated experiences, as when, for instance, exports, bank clearings, industrial failures, etc., are summated for periods of a day, a month, or a year. The ordinates represent facts realized only at the termination of, and not the characteristics of phenomena in, periods, deviations from which might be positive or negative. Under such circumstances, lines connecting successive ordinates are as much without meaning as lines connecting successive ordinates in discrete frequency series.

[1] On the purposes and methods of showing graphically ratio rather than absolute differences, see Fisher, Irving — "The 'Ratio' Method of Plotting Statistics," in *Quarterly Publications of the American Statistical Association*, June, 1917, pp. 577–601.

[2] On the purposes and methods of measuring price changes, see Chapter IX, *infra*.

They emphasize the direction or trend, but do not show the distribution of a fact at all possible periods of time over the range chosen. By no stretch of the imagination could it be assumed from the graphic representation alone whether the rates at which increments have been added to successive amounts within a given period were constant and uniform, or widely dissimilar. Measurements on successive ordinates must be made from the base line and not from the tops of preceding ordinates. The difference shown by the latter method merely reflects an excess or deficiency over past or future activity, as the case might be. Such series are discrete in a very definite sense, and the curves formed by joining successive ordinates have no other function than to aid the eye in judging direction or trend.

If fluctuations are violent and it is difficult to estimate the general direction or trend either for shorter or for longer periods, smoothing may be resorted to, but always with the understanding that its sole function is to clarify the movement and not to describe an ideal distribution. When it is done, it must follow the principles discussed below.

On the other hand, certain historical series represent, not accumulated facts at the close of arbitrary periods, but characteristic facts, deviations for the periods chosen being positive or negative, and coincident with the passage of time. Of such a nature are the curves describing, for arbitrary periods, changes in temperature, barometric pressure, expansion and contraction of metals under conditions of heat and cold, etc. For such series, ordinates should be erected at the middle points of the time-units and the tops connected either by straight, or preferably by smoothed, lines. In reality, such historical series are each composed of a succession of continuous frequency series, to which the rules and principles respecting smoothing are as applicable

as to continuous frequency series alone. Under such circumstances smoothed curves do far more than give a direction of trend. They more accurately describe the distribution at the individual periods and over the whole range than do the arbitrary measurements.

When related series are plotted on the same sheet, they should be designated by similar markings. Lines which lie closely together or frequently cross each other should be distinguished by dissimilar markings. Since color schemes are generally prohibitive, it is wise to make the choice of markings varied where many curves are drawn upon one sheet. Lines should always be broad enough to be readily followed, but not to sacrifice the accuracy of the ordinate unit.

(3) Purposes and Methods of Smoothing Historigrams

The methods of smoothing historigrams are subsidiary to the purposes to be accomplished by smoothing. If nothing better than a knowledge of general direction is desired, often one may rely wholly upon the free hand method. Smoothing in this manner is generally inaccurately done, however, and when averages or other summary expressions are read from smoothed curves, appreciable error results. If more exact knowledge is desired than that attainable by the rough method, and the series is cyclic in character, the method of "moving averages" or "progressive means" may be used. This method consists in plotting the averages (arithmetic) of the frequencies for the periodic cycles opposite the middle points (years or other time units), if the period chosen is of an odd number, or halfway between the two middle points if the number is even. This process is repeated throughout the entire series, each average plotted

being the result of lopping off one period and adding on another. By this process, the beginning and end of the period covered are not smoothed, but if the direction of the smoothed curve is definite, these portions may be completed, if it is thought necessary, by projecting the curve on the basis of the direction taken, or by assuming that the data, for a period long enough to complete the smoothed figure, have repeated themselves, or that the rate of increase or decrease has remained constant.[1]

This method, of course, is restricted to series in which cyclic or periodic changes are present. Care should always be taken to use periods which accurately coincide with a complete cycle. If, for instance, a period were used which corresponded to a half cycle, the resulting curve, while it would smooth out the minor fluctuations of the incomplete periods, would not materially affect the longer changes. If a period somewhat shorter or longer were taken, the smoothed curve would partake of the qualities of both the short- and long-time fluctuations. Its direction and significance would largely be indeterminate. Often no single period can be found which will accurately fit the cycles. They may not all be of the same length nor of the same magnitude. In cases where periods are so dissimilar that a distorted curve would result from using an average period, it is best not to employ the moving average method. Free-hand smoothing may then be used, but in any case the resulting curve must be interpreted in terms of the data and of the purpose or purposes for which it is smoothed.

If a series, although being historical, is of the discrete type, — the frequencies representing cumulations assignable to various periods, and the rate at which the increments are added being unknown, — a smoothed curve, made either

[1] See *infra*, Chapter XII, where this method is employed.

free hand, or according to the method of moving averages, represents nothing more than a series of approximations (averages) to the absolute quantities assigned to the periods treated. The smoothed curve can in no sense be looked upon as an accurate characterization of a series, the true or "normal" order of which has been distorted because of the units in which expressed. Long- or short-time fluctuations may be removed, but the fact remains that the measurements are discrete, and it is necessary to keep this in mind when interpreting the smoothed curve.

On the other hand, when an historical series of the second type — that in which frequencies although stated historically are typical of the period, or which, as is the case of temperature curves, record the exact condition currently — is smoothed, either by the rough and ready method or by moving averages, the resultant curve, while affected by extremes, probably more accurately characterizes the series, at least as theoretically distributed, than any unsmoothed curve could possibly do.

When historical series are compared with the purpose of correlating increase or decrease in one with increase or decrease, or the reverse, in the other, it is often necessary to reduce the short- and long-time fluctuations to mathematical bases and to treat them differently as purposes differ. This phase of the problem is treated in Chapter XII.

2. *Plotting Cumulative Historical Series* [1]

Historical series of the discrete type may be cumulated by successively grouping the frequencies at various unit-periods. In this respect they are not different from frequency series of the discrete type. The discussion of the interpreta-

[1] See Plate 18, Chapter VIII, and discussion.

tion given to the latter applies with equal force to the former. No significance, except that of judging the successive additions or subtractions, as the case may be, — depending whether the curve is read "up to and including" or "after and including," — can be attributed to the heights of the successive ordinates. No meaning can be attached to the lines connecting them, whether straight or smoothed, except as indicating general direction of change as cumulation proceeds. As is the case with all discrete series, whether simple or cumulative and whether frequency or historical, the lines connecting successive ordinates must be regarded only as aids to the eye and not as characterizations of ideal distributions.

For historical series of this type and treated in this manner, smoothing is not generally necessary, and when employed often has a tendency to smother the truth and to suggest license in the use of graphic methods. Neither should be cultivated.

IV. Conclusion

Both diagrammatic and graphic presentation of statistical data rightly viewed constitutes the art of statistical expression. Neither is necessary, although both are significant as preliminary to comparison, — the goal of statistical studies. The aim in this chapter has been to call attention to the most important considerations bearing upon the *science* connected with both, and not to the infinite uses which they may legitimately and illegitimately serve. It is their appeal, their smug finality, which suggest their virtues and at the same time conceal their weaknesses. Our purpose has not been to detract from their function, nor to agitate against their use, but solely to point out the cautions and conditions which make their employment scientific and their position

secure. This much, it is felt, it is necessary to do in view of the marked tendency to popularize them and to regard them as ends.

References

Bowley, A. L. — *Elements of Statistics*, Ch. VII, Sections I, II, III, IV, pp. 143–188.

Brinton, W. C. — *Graphic Methods for Presenting Facts*, Chs. IX, X, pp. 149–163, 164–199, respectively.

Elderton, W. P. and Ethel M. — *Primer of Statistics*, Ch. III, pp. 23–39.

King, W. I. — *Elements of Statistical Methods*, Ch. XI, pp. 97–120.

Thorndike, E. L. — *An Introduction to the Theory of Mental and Social Measurements*, Ch. III, pp. 28–41.

Yule, G. U. — *An Introduction to the Theory of Statistics*, Ch. VI, pp. 75–105.

Marshall, Alfred — "On the Graphic Method of Statistics" in the *Jubilee Volume, Journal of the Royal Statistical Society*, 1885.

Fisher, Irving — "The 'Ratio' Chart for Plotting Statistics," in *Quarterly Publications of the American Statistical Association*, June, 1917, pp. 577–601.

CHAPTER VIII

AVERAGES AS TYPES

I. Introduction — General Statement

The progress of the treatment has carried us toward a single goal — that of comparison. Step by step the conditions and limitations which must be imposed in the collection of primary and in the use of secondary data have been considered. In their various aspects, the collection and classification of data and the devices currently in use and advocated for use in diagrammatic and graphic presentation have been discussed. The limits of the latter have been emphasized and the purposes and the consequences of the former considered. Throughout all stages of the treatment the limitations of statistical method, when used alone, have been acknowledged and emphasis placed particularly upon the difficulties of reducing to numerical bases the vital considerations connected with economic phenomena. The complexity of economic problems, and the many angles from which they must be considered before weight can be attributed to conclusions based alone upon statistical data, ought to stand out distinctly as one of the net results of the discussion.

If the collection, classification, and arrangement of statistical data present problems, — *i.e.* if the *processes* involved offer difficulties, — how much more serious must be the problems when, in order to explain, describe, or establish the causal relationships between phenomena, the *results* or con-

clusions arising out of the use of these processes become the tools with which we operate. It is then not only necessary that the conditions surrounding enumeration, observation, and summarization of statistical data be appropriate, but that the conclusions which are deduced from them be logically sound and properly employed! And yet, statistically, comparison is the end toward which all previous steps are but preparatory.

The data of economics are highly complex. They relate to conditions, evidences of which are not reducible to absolute uniformity of expression. They exhibit themselves in varying and changing proportions. Economic phenomena exist as cause and effect of other phenomena, and not independently. They must be dealt with as related forces. If they are inherently complex, so likewise are the methods by which they are described or measured. Simple units will not often suffice. Definitions are difficult to formulate, and to adhere to them strictly in all stages of work is frequently impossible. Care, judgment, insight, and caution are eternally necessary to guard against mistaken views, the assignment of cause for effect, the omission of qualifying or significant facts, the formation of false judgments, etc.

For the focusing of judgment which comparison requires, concentrated or summary expressions are necessary. We seek for units of analysis here as we sought for units of collection earlier. Data in all their inclusiveness and in all their detail cannot readily be compared as between periods, times, or conditions. Some single expression which gathers into itself all the significant characteristics of complex data is required. We seek in actual life for an average performance, an average load, an average student or clerk, an average day, an average market, average conditions, etc., in order to bring things into relation. In general discussions such con-

cepts are used loosely, and frequently important matters are settled by employing no more definite or restricted terms. The willingness to be content with a single expression as a substitute for complex detail is often an evidence of ignorance either of the difficulties in making comparison or of the limitations of summarizing expressions.[1] Invariably to speak and write of economic problems in terms of averages connotes a willingness either to be content with *general notions* — often so general as to be meaningless — or indifferently to employ tabloid expressions as accurate characterizations of complex things. Short cuts to the goal of comparison are too often preferred to the circuitous but more certain paths. Attempts are too frequently made to compare or contrast economic phenomena by appeal to averages in the form of median, mode, or arithmetic mean, where in reality not only are comparisons invalid but the data themselves do not admit of so being summarized. That fundamental canon which cautions against relating things to conditions incapable of producing them is flagrantly violated, and assurance of correct thinking found in the belief that we are dealing with "average conditions." This complacent belief may suffice to lull the ignorant into a state of blind indifference, but to those who are unwilling to allow themselves thus to be beguiled it offers little guaranty of intellectual repose.

Rarely, if ever, does a summary expression carry with it the same amount of truth as do detailed data.[2] An average often suffices to give one a more convenient and more easily

[1] Watkins speaks of averages as "representative numbers" and as containing "the gist, if not the substance, of statistics." G. P. Watkins, "Theory of Statistical Tabulation," *Publication of the American Statistical Association*, December, 1915, p. 752.

[2] Venn, Dr. John, "On the Nature and Use of Averages," *Journal of the Royal Statistical Society* (London), Vol. LIV, 1891, pp. 429–448, at page 433.

grasped view of a difficult and complex situation than do detail, but its seeming oneness and finality are the precise sources of its limitations. The same numerical average may be computed from widely different detail. Yet it may be these in which interest lies. These, of course, are sacrificed in case reliance is placed alone in averages. An average in statistical methods has an analogous function to that of a generalization in inductive logic, viz., as a means of crystallizing into a single expression or of formulating into a single concept a general truth. As experimentation and observation precede the formulation of a general truth in logic, so in statistics does analysis of numerical detail precede their summarization into a single expression. The use of an average presupposes a knowledge of the data out of which it grows, a clear conception of the peculiar features of the particular average used, and a certain mastery of the whole subject treated so as to be sure of the validity of the comparison which its use involves.

The pertinency of the discussion will probably be more apparent as we treat descriptively and functionally the more general forms of averages in current use, their particular merits for different kinds of data, and the methods of computing them.

II. Averages Descriptively Considered

The types of averages here dealt with are those in common use. At this stage of the discussion, a simple definition of each kind will suffice. The peculiar qualities will develop later. It may then be necessary to redefine them in terms of all their uses and implications.

The *arithmetic mean* or *average* of a series is that amount which is derived by dividing the sum or aggregate of the parts by the number. It is solely a numerical concept.

The *median* of a series is that item — actual or estimated — in a series, when arranged consecutively, which divides the distribution into equal parts. When the number of items is even, it is halfway between the two middle terms; when the number is odd, it is the middle term. Like an arithmetic average or mean, it is primarily a numerical expression.

The *mode* of a series is that item or term which is most characteristic or common. It represents the typical fact and always relates to a condition which is actually represented, — thus not being restricted simply to a numerical concept.

On the basis of the methods by which these averages are computed the following classes may be distinguished: (1) averages requiring all of the data for their computation; (2) averages requiring only a part of the data for their computation; (3) averages which of necessity are represented in a series; (4) averages which by chance may be so represented but are primarily numerical concepts; (5) averages which are affected equally by both size and number of the items measured; and (6) averages in which the frequency must be known but in which it is necessary to know only the approximate size of the units to which the frequencies apply. The arithmetic mean clearly falls in the first class since the entire aggregate is included. The median and the mode fall in the second class. In class three any one of the averages may fall but the mode is always included. In class four belong the median and the arithmetic mean. In class five falls the arithmetic mean, and in class six both the median and the mode. The precise reasons for, and significance of this classification will be seen in the discussion of these averages.

III. The Arithmetic Mean or Average

1. *What the Arithmetic Mean or Average Is*

The arithmetic mean or average is undoubtedly the most familiar average in current use. Indeed, it is the only one customarily employed in elementary studies, and an explanation of it might seem unnecessary. It is the one commonly used in the ordinary transactions of business and commercial life, and for this reason possesses certain value. It represents the center of gravity or balancing point of a group or a number of items, the differences or deviations in excess being exactly counterbalanced by the difference or deviations in defect of it. In its computation all items are considered, each being given an importance equivalent to its size and its distance above or below the average. It is primarily a mathematical concept, and is susceptible of duplication from great varieties of distributions. In this fact lies its weakness when it is used as a substitute for a complete description of a series.

To be specific. The arithmetic average of the series 8, 9, 10, 11, 12, 13, 14, is 11. Likewise the arithmetic average of the series 8, 8, 8, 9, 9, 9, 10, 10, 10, 11, 11, 11, 12, 12, 12, 13, 13, 13, 14, 14, 14, is 11. The same is true of the following series 2 and 20; 9, 9, 4, 22; 3, 1, 1, 1, 99, 1, 1, 1, 1, 1, 11, and almost any number of other combinations which one might choose. When an average is thus so wholly independent of the order of the series, the number of items, and of their relative size, it has serious limitations for all uses in which the character of a distribution is of vital concern. Moreover, the arithmetic mean may really never be represented in a series, as for instance, when 2 and 20, or 9, 9, 4 and 22 are averaged. Nothing typical is thus revealed.

The only thing which we have is a mathematical expression of an aggregate divided by a number of items. It is clear that this form of average has serious limitations when applied to widely different conditions, or to data describing them, and must be used with extreme caution. Especially is this true when it does not represent an actual fact. An arithmetic mean wage-rate, computed for a group of employees, may fail to describe a single actual rate. It may also be so different from those that are characteristic as to lead to ridiculous conclusions if reliance is placed in it. The inclusion of a single exceptional rate might invalidate its use. Instances will arise, of course, when the exceptional circumstance or fact should be included. The thing which is now sought to be emphasized is that an arithmetic mean *per se* gives no guaranty of the distribution or nature of the items which make it up. It is, therefore, a crystallizing or summating expression to be used with extreme care in all series in which distributions are irregular and in which items are noticeably dissimilar. When used it should always be accompanied by some other forms of summary expressions where there is any question as to its legitimacy.

In mathematical science the position of the arithmetic mean or average is clearly established. One authority, in speaking of its value in connection with *Adjustment of Observations*, says, "If we have n observed values of an unknown, all equally good so far as we know, the most plausible value of the unknown (best value on the whole) is the arithmetic mean of the observed values."[1] Speaking further, the same authority says "when the number of observed values is very great, the arithmetic mean is the true value."[2] This fact is based upon the principle that in the absence of

[1] Wright, T. W., and Hayford, J. H., *The Adjustment of Observations*, p. 10.
[2] *Ibid.*, p. 11.

bias, large errors are less frequently encountered than small ones, and that they tend to be distributed about a true value according to the laws of chance. That is, positive and negative errors of the same size occur with the same frequency. The fact that measurements of great mathematical accuracy, or subject to pure chance selection, are rarely found in economics and in business affairs robs this average of much of its mathematical importance.[1] Too frequently, observations are not all "equally good," and do not fall into symmetrical and continuous series. Too often they are vitally affected by limitations of the units, by the bias of the collector or of those who supply them, and frequently do not admit of accurate measurement.

The function and peculiarities of the arithmetic average may further be illustrated by a discussion of the means of its computation. At the same time the difference between simple and weighted averages will be developed.[2]

2. *How the Arithmetic Mean is Computed*

As noted above, the arithmetic mean is the center of gravity of a distribution. This fact may conveniently be illustrated by the use of an imaginary rod upon which certain weights are suspended at intervals. If it is desired to determine the arithmetic mean wage-rate of the following distribution, it may of course be done by summating or totaling the rates and dividing by the number of instances.

[1] Certain mathematical properties of the arithmetic mean are discussed in Yule, G. U., *An Introduction to the Theory of Statistics*, pp. 114 ff. and in Wright and Hayford, *op. cit.*, Chap. I.

[2] The methods and significance of weighting are further discussed in Chapter IX.

TABLE A

TABLE SHOWING WAGE-RATES AS BASES FOR THE COMPUTATION OF A SIMPLE ARITHMETIC MEAN RATE

THE UNIT OR AMOUNT AVERAGED	THE NUMBER OF TIMES EACH UNIT IS ENCOUNTERED (The Weight)
$39.00	9
$2.00	1
4.00	1
3.00	1
6.00	1
3.00	1
8.00	1
5.00	1
3.50	1
4.50	1

$39.00 divided by 9 = $4.33 = the arithmetic mean or average. That is, if upon an extended rod properly scaled equal weights (in this case one) be suspended at the measurements here shown, the rod will balance at the point $4.33. This condition is diagrammatically illustrated by Figure A, Plate 16. On the other hand, if we use the same units and assign to them representation greater than unity, but at the same time retain the same proportion between the weights (*i.e.* the frequency with which they occur), the average will not be changed. The rod will balance at the same point. Diagrammatically, this adjustment is illustrated in Figure B, Plate 16.

If weights are greater than unity and their positions on the scale are altered, the resulting average will be different. If the adjustment has been according to chance the difference, however, will be small. From this would follow the conclusion that if data are accurately chosen and are representa-

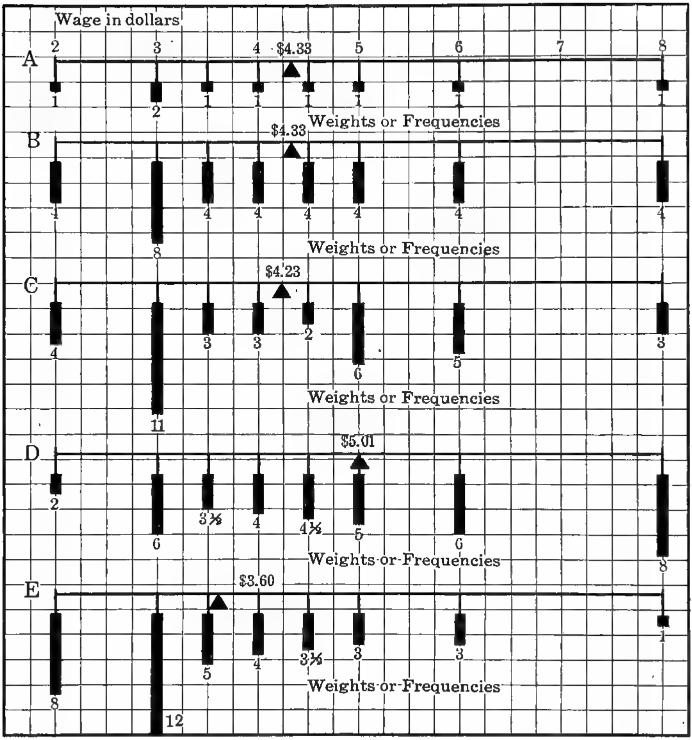


PLATE 16

Diagrams Illustrating the Nature of the Arithmetic Mean when Items are Differently Weighted.

tive, weights may largely be ignored. Of course, to say that data are thoroughly representative is only one way of saying that weights have been properly distributed. Taking the same units as above, and the following chance weights, the average is reduced by only \$.10, notwithstanding the fact that the difference between the extreme weights is 7, and that a weight of one item is $4\frac{1}{2}$ times as large as that of another.

TABLE B

Table Showing Wage-rates with Number of Persons Receiving Them as Bases for Computing a Weighted Arithmetic Mean Rate

The Unit or Amount Averaged	The Number of Times Each Unit is Encountered (The Weight)	Product of the Weight Times the Unit
Total	37	\$156.50
\$2.00	4	8.00
4.00	3	12.00
3.00	9	27.00
6.00	5	30.00
3.00	2	6.00
8.00	3	24.00
5.00	6	30.00
3.50	3	10.50
4.50	2	9.00

The resulting average is \$156.50 — the aggregate — divided by the number of items — the sum of the weights — and equals \$4.23. Diagrammatically, this combination of weights and units is shown in Figure C, Plate 16.

By arbitrarily adjusting the weights or frequencies with which each item is repeated, the average may be increased or decreased at will. In column 1 below, the weights have

been chosen in such a manner that the items larger than the average (when all items are taken once) are given heavy weights and those below the average light weights, the amount of importance varying directly with the size of the unit. In column 2 the reverse order of weights is chosen. Diagrammatically, the effect of both processes is shown in Figures D and E, respectively, Plate 16.

TABLE C

TABLE SHOWING WAGE-RATES WITH NUMBER OF PERSONS RECEIVING THEM AS BASIS FOR COMPUTING WEIGHTED ARITHMETIC MEAN RATES

THE UNIT OR AMOUNT AVERAGED	COL. 1 THE NUMBER OF TIMES EACH UNIT IS ENCOUNTERED (THE WEIGHTS)	PRODUCTS OF UNITS AND WEIGHTS	COL. 2 THE NUMBER OF TIMES EACH UNIT IS ENCOUNTERED (THE WEIGHTS)	PRODUCTS OF UNITS AND WEIGHTS
Total . . .	39	$195.50	39.5	$142.25
$2.00	2	4.00	8	16.00
4.00	4	16.00	4	16.00
3.00	3	9.00	6	18.00
6.00	6	36.00	3	18.00
3.00	3	9.00	6	18.00
8.00	8	64.00	1	8.00
5.00	5	25.00	3	15.00
3.50	$3\frac{1}{2}$	12.25	5	17.50
4.50	$4\frac{1}{2}$	20.25	$3\frac{1}{2}$	15.75
Average		5.01		3.60

By thus arbitrarily adjusting the weights, the exact sizes being essentially within the limits of those assigned by chance, the resulting average has been increased in the first case (column 1) over that arrived at by assigning equal weights, by $.68, and over that gotten by assigning chance weights, by $.78. In the second case, the average as compared to

that obtained by using equal weights has been decreased \$.73, and when compared to that received by using chance weights \$.63. The difference obtained by arbitrarily shifting the weights is \$1.41 as compared to \$.10 when equal and chance weights are used. The interesting fact is suggested that this average is a function of the weights that are used, tending to be larger than the simple average of an unweighted or equally weighted series when items larger than it are heavily weighted, and smaller than it when smaller items are heavily weighted.

Weights should always be carefully chosen and the validity of weighting clearly established. When weights are chosen at random, the resulting average is affected very little by their absolute size. The more nearly bias is eliminated, the closer will the weighted approach the simple average. By taking the distribution of wage-rates above and assigning to them pure chance weights (done by drawing by chance from a group of numbers marked with figures from 1 to 29, both inclusive) the following averages in four trials were determined: \$4.43, \$4.26, \$4.29, and \$4.04 — average \$4.27, which agrees closely with the simple average.[1]

[1] The following are chance weights used in this experiment:

Units	1st Trial	2d Trial	3d Trial	4th Trial
\$2.00	25	22	13	23
4.00	22	24	21	14
3.00	17	11	23	6
6.00	23	26	24	28
3.00	1	27	14	15
8.00	15	16	10	1
5.00	27	16	20	10
3.50	12	25	19	2
4.50	21	23	24	3

(The student is advised to try others.)

The computation of an arithmetic mean is generally readily done by the ordinary method. In some instances, however, particularly where frequency groups are dealt with, it is easier to proceed in a different manner. On the principle that the sum of the deviations from the true average, signs considered, equals zero, an average may be assumed as a starting point, the deviations calculated and corrected for error, and the true average determined. The use of this method for various arrangements of data may be illustrated as follows. It is desired to calculate the simple average wage-rate of the following distribution. Assume as a trial average \$5. The sum of the minus deviations = −\$10; the sum of the plus deviations = \$4; the algebraic sum = −\$6. This must be divided by the sum of the frequencies, 9, and added to \$5.

$$\frac{-\$6}{9} = -\$.67. \quad \$5.00 + (-\$.67) = \$4.33, \text{ the true average.}$$

TABLE D

TABLE GIVING DATA FOR COMPUTING THE ARITHMETIC MEAN BY THE "SHORT-CUT" METHOD

UNITS OR AMOUNTS	FREQUENCIES	DEVIATIONS		NET DEVIATIONS
		−	+	
Total	9	\$10.00	\$4.00	−\$6.00
\$2.00	1	\$3.00		
4.00	1	1.00		
3.00	1	2.00		
6.00	1		1.00	
3.00	1	2.00		
8.00	1		3.00	
5.00	1			
3.50	1	1.50		
4.50	1	.50		

If frequencies are larger than unity, the method is not changed. The only necessary step is to multiply the deviations by their respective frequencies. Thus:

TABLE E

TABLE GIVING DATA FOR COMPUTING THE ARITHMETIC MEAN BY THE "SHORT-CUT" METHOD

UNITS OR AMOUNTS	FREQUENCIES	DEVIATIONS		DEVIATIONS TIMES THE FREQUENCIES		TOTAL NET DEVIATIONS
		−	+	−	+	
Total	163			$161.50	$68.00	−$93.50
$2.00	25	$3.00		75.00		
4.00	22	1.00		22.00		
3.00	17	2.00		34.00		
6.00	23		$1.00		23.00	
3.00	1	2.00		2.00		
8.00	15		3.00		45.00	
5.00	27					
3.50	12	1.50		18.00		
4.50	21	.50		10.50		

$$-\$93.50 \div 163 = -\$.57.$$

$$\$5.00 + (-\$.57) = \$4.43 = \text{the arithmetic mean.}$$

When dealing with frequency groups, the actual distribution of the items within the groups is not known, and it is necessary to multiply them by some characteristic term. Except when groups are very wide or data are distinctly of the discrete type, it is admissible to consider the numbers

within the groups to be uniformly dispersed and to multiply the frequencies by the middle terms. Taking the following frequency distribution of wage-rates, the arithmetic mean may be calculated both by the regular and short-cut methods.

TABLE F

TABLE GIVING DATA FOR COMPUTING AN ARITHMETIC MEAN FROM FREQUENCY GROUPS

UNITS OR AMOUNTS	FREQUENCIES	PRODUCTS OF FREQUENCIES AND THE UNITS (Middle Terms)
Total	434	$3,923.00
$5.00 to $5.99	15	82.50
6.00 to 6.99	40	260.00
7.00 to 7.99	66	495.00
8.00 to 8.99	91	773.50
9.00 to 9.99	113	1,073.50
10.00 to 10.99	49	514.50
11.00 to 11.99	30	345.00
12.00 to 12.99	27	337.50
13.00 to 13.99	2	27.00
14.00 to 14.99	1	14.50

$3,923 ÷ 434 = $9.04 = arithmetic mean or average.

If we proceed by the method of computing the deviations from an *assumed* average, the steps are not different from those used above when the data were not arranged in groups, except that it is necessary, as in the case immediately above, to assume a uniform distribution throughout each group. The method is shown in the following example, using the

data immediately above. The assumed average is $9.50, *i.e.* the item halfway through the group, $9.00 to $9.99.

TABLE G

TABLE GIVING DATA FOR COMPUTING AN ARITHMETIC MEAN BY THE "SHORT-CUT" METHOD FOR FREQUENCY GROUPS FROM AN ASSUMED AVERAGE

UNITS OR AMOUNTS	FREQUENCIES	DEVIATIONS FROM THE ASSUMED AVERAGE, $9.50		PRODUCTS OF DEVIATIONS AND FREQUENCIES		NET DEVIATIONS
		−	+	−	+	
Total . .	434			$403.00	$203.00	− $200.00
$5.00 to $5.99	15	$4.00		60.00		
6.00 to 6.99	40	3.00		120.00		
7.00 to 7.99	66	2.00		132.00		
8.00 to 8.99	91	1.00		91.00		
9.00 to 9.99	113					
10.00 to 10.99	49		$1.00		49.00	
11.00 to 11.99	30		2.00		60.00	
12.00 to 12.99	27		3.00		81.00	
13.00 to 13.99	2		4.00		8.00	
14.00 to 14.99	1		5.00		5.00	

− $200 ÷ 434 = − $.46. That is, the net average deviation does not equal zero, but − $.46. Therefore, in order to determine the true average (from which the sum of the deviations equals zero) it is necessary to add − $.46 to the assumed average, $9.50, thus giving $9.04 as the true average. The plus and minus deviations, calculated in the same manner from the *true average*, $9.04, are given below.

TABLE H

Table Showing the Effect of Computing the Arithmetic Mean from the True Average for Data in Frequency Groups

Units or Amounts	Frequencies	Deviations from the True Average, $9.04		Products of Deviations and Frequencies		Net Deviations
		−	+	−	+	
Total . . .	434			$305.48	$305.12	− $.36 [1]
$5.00 to $5.99	15	$3.54		53.10		
6.00 to 6.99	40	2.54		101.60		
7.00 to 7.99	66	1.54		101.64		
8.00 to 8.99	91	.54		49.14		
9.00 to 9.99	113		$ .46		51.98	
10.00 to 10.99	49		1.46		71.54	
11.00 to 11.99	30		2.46		73.80	
12.00 to 12.99	27		3.46		93.42	
13.00 to 13.99	2		4.46		8.92	
14.00 to 14.99	1		5.46		5.46	

When frequency groups are all of equal size it is often a saving of time to compute the deviations from an assumed average in terms of the "steps" which successive groups are above or below the group containing the assumed average, and later to convert the net "step-deviations" back into real deviations by multiplying by 1, in case the step is unity, or by 2 in case it is two or by $\frac{1}{2}$ in case it is one half, etc. Using the above distribution, but assuming a different average, we have computed the arithmetic mean by the "step" method.

[1] This negligible difference is due to the fact of taking the average at $9.04. The exact average is $9.039 +.

TABLE I

TABLE GIVING DATA FOR COMPUTING THE ARITHMETIC MEAN BY THE "STEP-DEVIATION" METHOD FOR FREQUENCY GROUPS FROM AN ASSUMED AVERAGE

Units or Amounts	Frequencies	Step-Deviations from the Assumed Average, $12.50		Products of "Steps" and Frequencies		Net "Step" Deviations
		−	+	−	+	
Total . . .	434			1,506	4	−1,502
$5.00 to $ 5.99	15	7		105		
6.00 to 6.99	40	6		240		
7.00 to 7.99	66	5		330		
8.00 to 8.99	91	4		364		
9.00 to 9.99	113	3		339		
10.00 to 10.99	49	2		98		
11.00 to 11.99	30	1		30		
12.00 to 12.99	27					
13.00 to 13.99	2		1		2	
14.00 to 14.99	1		2		2	

$-1502 \div 434 = -3.46$. $-3.46 \times \$1.00$ (the size of the group) $= -\$3.46$. $12.50 (the assumed average) $+ (-\$3.46) = \$9.04$ = the true average.

Where groups are not uniform in size, this method cannot be employed without considerable difficulty. When they are uniform, however, much trouble in multiplying is avoided by computing the deviations in round numbers and subsequently by converting them back into terms of the size of the "step." The following table illustrates the method when groups are of unequal size.[1] In such cases it is far simpler to proceed in the regular manner by multiplying through in the first instance.

[1] This method involves "averaging averages" and is of doubtful value.

TABLE J

TABLE GIVING DATA FOR COMPUTING THE ARITHMETIC MEAN BY THE "STEP-DEVIATION" METHOD FROM AN ASSUMED AVERAGE WHEN THE GROUPS ARE OF UNEQUAL SIZE [1]

GROUPS			FREQUENCIES	"STEP-DEVIATIONS"		PRODUCTS OF "STEPS" AND FREQUENCIES		NET "STEP-DEVIATIONS"
Size	Width	Center		−	+	−	+	
Total			30,454					
Total			24,885			13,976	15,242	+ 1266 [3]
[2] Less than 6¢	2	5	99	4		396		
6¢–8¢	2	7	661	3		1,983		
8¢–10¢	2	9	2,722	2		5,444		
10¢–12¢	2	11	6,153	1		6,153		
(1) 12¢–14¢	2	13	6,007					
14¢–16¢	2	15	4,926		1		4,926	
16¢–18¢	2	17	2,635		2		5,270	
18¢–20¢	2	19	1,682		3		5,046	
Total			5,076			2,604	468	– 2136 [4]
20¢–25¢	5	22.5	2,604	1		2,604		
(2) 25¢–30¢	5	27.5	2,004					
30¢–35¢	5	32.5	468		1		468	
Total			291					
(3) 35¢–45¢	10	40	291					[5]
Total			202			109	33	– 76 [6]
45¢–60¢	15	52.5	109	1		109		
(4) 60¢–75¢	15	67.5	60					
[2] 75¢ and over	15	82.5	33	1			33	

[1] Data taken from Report of the Tariff Board on Schedule "K," Vol. IV., Part 5. *House Doc. 342, 62d Congress, 2d Session*, p. 997.

Notes, [2], [3], [4], [5], and [6] on following page.

In summarizing the discussion of the arithmetic mean, attention should be called to the fact that it is easily understood, is readily calculated, is in everyday use, and is affected by all the items in a series. Indeed, when nothing more is wanted, as a summarizing expression, than the total divided by the sum of the parts, it thoroughly meets the need. But in statistical analysis of economic problems the needs generally run far beyond this. It is frequently the detail which is of most importance and which is so often concealed by the arithmetic mean. It is too susceptible to the extraordinary, too much affected by the exceptional, to serve all purposes equally well. Various checks may be imposed in order to test its validity for a definite purpose. The details themselves may be submitted. But this is often impossible, since the employment of an average is an indication of a desire or of a necessity to be free from detail. Other averages may be computed for purposes of comparison, and it is to a discussion of these to which we now turn.

[2] Width of group assumed to be the same as that of the class to which it belongs.

[3] + 1266 ÷ 24,885 = .0509. .0509 × 2¢ (the width of the group) = \$.001018. \$.13 + \$.001018 = \$.1310 (average of the first group).

[4] − 2136 ÷ 5076 = − .421. − .421 × 5¢ (the width of the group) = − \$.02105. \$.275 + (−\$.02105) = \$.254 (average of the second group).

[5] \$.40 is the average of the third group.

[6] − 76 ÷ 202 = − .376. − .376 × 15¢ (the width of the fourth group) = − \$ .05640. \$.675 + (− \$.05640) = \$.6186 (average of the fourth group).

Groups	Averages	Weights	Products of Weights and Averages
Total	\$.1573	30,454	\$4790.5962
(1)	\$.1310	24,885	3259.9350
(2)	.2540	5,076	1289.3040
(3)	.4000	291	116.4000
(4)	.6186	202	124.9572

IV. The Median

1. *What the Median Is*

The median has been defined as the item in a series, when arranged consecutively, which divides the distribution into equal parts. While it is generally called an average it is more accurately a measure of partition or distribution. It can be said to be characteristic of the other members of a series only in case they are uniformly dispersed around it. It divides frequencies into equal parts and not the units to which they apply. Indeed, the exact size of an item measured need not be known. The only thing necessary is to be able to place it in a distribution so that the order of arrangement is consecutive. Unlike the arithmetic mean, it is not primarily a mathematical concept, since it may be used where numerical significance is not attributed to the factors averaged, as, for instance, in the grading of pupils, salesmen, etc., simply by placing them in their order of excellence. This, of course, means nothing more than that relative rank is established. The middle position is then determinable. Yet it is like the arithmetic mean in the fact that the middle or median quantity itself need not be represented in a series. How accurately a distribution is characterized by the median alone depends almost entirely upon its nature. Perhaps we can get a clearer view of its meaning if we compute it for a variety of distributions. Remembering that it is that item which divides a series, consecutively arranged, into equal parts, and substituting n for the number of items in the series, the expression $\frac{n+1}{2}$ may be used as a basis for its computation.

2. *How the Median is Computed*

Using the data in Table A, p. 242, but rearranging the units in an ascending order (a thing unnecessary in the computation of the arithmetic mean), we get the following series:

TABLE K

TABLE GIVING DATA FOR COMPUTING THE MEDIAN

UNIT	FREQUENCIES
Total	9
\$2.00	1
3.00	1
3.00	1
3.50	1
4.00	1
4.50	1
5.00	1
6.00	1
8.00	1

Applying the formula, $\frac{n+1}{2}$, when $n = 9$, we get $\frac{9+1}{2} = 5$, *i.e.* the fifth item in the series divides it into equal parts. Counting down from the smallest item, or up from the largest one — a matter of indifference — \$4.00 is found to be the median. It should be noticed, however, that the thing which is really divided into two equal parts is the total frequency, and not the items to which the frequencies apply. That is, \$4.00 is only \$2.00 away from the first item, but \$4.00 away from the last. Moreover, the \$2.00 in

the series is of as much importance in determining the median as is $8.00. It is quite different, of course, respecting the arithmetic mean. Moreover, retaining the frequencies as above, every item in the series except the middle one may be changed — the only limitation being that the order must remain ascending — and the median remain the same. Let us arrange some changes in the form of a table, still leaving the median $4.00, and compute the corresponding arithmetic mean in each case.

TABLE L

TABLE GIVING DATA SHOWING THE EFFECT OF CHANGES OF DISTRIBUTION ON THE MEDIAN AND THE ARITHMETIC MEAN

FREQUENCIES	UNITS AND ILLUSTRATIONS					
Total 9	1st	2d	3d	4th	5th	6th
1	$2.00	$1.00	$3.99	$4.00	$ .25	$2.00
1	3.00	1.00	3.99	4.00	.50	3.00
1	3.00	1.00	3.99	4.00	.75	3.00
1	3.50	1.00	3.99	4.00	1.00	3.50
1	4.00	4.00	4.00	4.00	4.00	4.00
1	4.50	4.00	4.01	4.00	4.00	4.50
1	5.00	4.00	4.01	4.00	4.00	5.00
1	6.00	4.00	4.01	4.00	4.00	6.00
1	8.00	4.00	4.01	4.00	4.00	10,000.00
Median	4.00	4.00	4.00	4.00	4.00	4.00
Arith. Mean	4.33	2.67	4.00	4.00	2.50	1,114.45

The median is invariably the 5th item, all others being important exactly in proportion to their frequency but not according to their amount. The median retains its stability so long as the central item does not change. Hence it is a

desirable "partition expression," — average, — to use only when the central groups are of interest, or where a distribution is regular and uniform. The exact size of the extremes or of any single item, except the middle one, may be ignored, the only thing necessary being a knowledge of their frequency and position above or below the median. All frequencies might be identical and the median alone never reveal the fact. This is true also of the arithmetic mean of a series of uniform frequencies. The deviations in this case equal zero, but this is true of any combination of frequencies howsoever arranged or of whatever size.

When the number of items is even and the units to which the frequencies apply are not expressed in groups, — that is, when the exact and not the approximate sizes are stated, — the median is arbitrarily taken as half-way between the two middle items. Of course, this assignment is purely arbitrary, and for all series other than those that are continuous, *i.e.* those in which the measures given are in reality only approximations of the true measures, and in which the differences would shade into each other by imperceptible gradations, if the number of separate measures were vastly increased — it should be considered as approximate. The exact median in this case is hardly more independent than when the number of items is odd. It is now determined not by one term, but by two, and these may be much alike or widely different. This is evident by an examination of the table immediately above. If to *Illustration* 1 the item $2.00 is added, the median is $3.75, *i.e.* it lies half-way between $3.50 and $4.00. If to *Illustration* 2 the item $8.00 is added, the median is still $4.00, and will continue to be $4.00 until more than 8 additional items are added, the only limitation being that they must be more than $4.00, but they may be any amount more. If to the series in *Illustration* 2, — $1.00, $1.00, $1.00,

\$1.00, \$4.00, \$4.00, \$4.00, \$4.00, \$4.00, — one item of each of the following is added: \$600.00, \$10,000.00, \$12,999.99, \$13,000.00, and \$14,621.47, the median is still \$4.00 as in the case without these exceptional numbers. The arithmetic average, however, is changed from \$2.67 to \$3,660.39.

In dealing with discrete series, one should rarely attempt to compute exact medians. Too great accuracy may result in making this average nothing more than a mathematical concept, ill suited to the units in which the data are expressed, and one wholly determined by the relation of the two middle terms. In continuous series the problem is different, inasmuch as the data used are generally samples and serve only more or less imperfectly to describe an ideal distribution. The median, of course, may be used in discrete series, but care should be taken not to assign too definite a position to it by refined methods of interpolation.

When data are arranged in frequency groups, the problem of determining the median is the same as it is when they are not grouped, except that it is necessary arbitrarily to distribute the frequencies within the groups in order to interpolate for the exact median. What is wanted is to locate not only the *median group*, but the *median item* in the group, in order to divide the series in half. To write the units in groups, assigning a frequency to them thus approximately measured, rather than to write them individually with the corresponding frequencies, makes it necessary to approximate the items within groups. When groups are small, in the case of discrete series, or when distributions are of the continuous type, the assumption of a uniform distribution is sufficiently accurate for most purposes. The error is not a seriously disturbing factor. The method by which the median of a series arranged in frequency groups is determined is illustrated in the following example, using the data from Table F.

TABLE M

TABLE GIVING FREQUENCY DATA FOR THE COMPUTATION OF THE MEDIAN

UNITS OR AMOUNTS	FREQUENCIES
Total	434
\$ 5.00 to \$ 5.99	15
6.00 to 6.99	40
7.00 to 7.99	66
8.00 to 8.99	91
9.00 to 9.99	113
10.00 to 10.99	49
11.00 to 11.99	30
12.00 to 12.99	27
13.00 to 13.99	2
14.00 to 14.99	1

In this instance the n in the formula is 434. By writing frequencies in this form, the necessity is obviated of listing each separate item, falling within the groups the number of times it appears. In determining the arithmetic mean in Table F, the frequencies were multiplied by the respective middle terms, on the assumption that the items through the groups were uniformly dispersed. Making the same assumption here, the median group is calculated by the formula $\frac{n+1}{2}$. $n = 434$. $\frac{434+1}{2} = 217\frac{1}{2}$, that is, the group containing the $217\frac{1}{2}$ item — wage-rate in this case — is the median group. Counting down from the smallest item, the group \$9.00 to \$9.99 is found to contain all the items between 212 and 325. The $217\frac{1}{2}$ man's wage-rate is, therefore, located within this group. On the assumption that the 113 men

whose wage-rates fall within the group \$9.00 to \$9.99 (inclusive) are uniformly distributed in the order of the size of their rates, the wage-rate which is half-way between that received by the 217 and the 218 man — that is, the $5\frac{1}{2}$ man in the group — is $\frac{5\frac{1}{2}}{113} \times \$1.00$ greater than \$9.00, *i.e.* the position of the first man in this group.[1] This gives him a wage-rate of \$9.049, which corresponds very closely to the arithmetic average, \$9.04.

In this example we are dealing with wage-rates — a discrete series — and the median is stated with sufficient accuracy when assigned to the lowest quarter of the group \$9.00 to \$9.99 — say, ± \$9.15. Weekly rates are not normally quoted in smaller units than quarter dollars, and it is inadvisable to strive for too great accuracy of expression. The degree of precision with which the median is determined largely depends upon the character of the dis-

[1] In order to have the 113 men distributed throughout this group uniformly and to have the same apply to the groups immediately following and preceding, it would be impossible to assign a man to the last unit of a preceding group and to the first unit of the succeeding group. To do this would result in a concentration at this point. Zizek, in discussing an analogous point, says: "We can distribute 10 values in a class of 200 cents breadth so that the first and the last values coincide with the limiting values of the class; so that the first item coincides with the inferior limit while the last value is as far distant from the superior limit as are the items from each other; or, so that the last item coincides with the superior limit while the first item is as far distant from the inferior limit as are the items from each other. None of these three distributions seems to be free from objection. The first kind of distribution, if carried out in the adjoining classes, would give two items at each class limit. The second and third kinds of distribution do not correspond at all to the postulate of a uniform distribution within the classes. The most correct way of distributing the items uniformly is to assume that they occur at equal intervals even when this distribution is extended to the adjoining classes. To fulfill this condition the first and last of the items belonging to the class must be removed from the class limits to a distance which corresponds to half the magnitude of the interval existing between the items belonging to the class." *Statistical Averages*, pp. 208–209.

tribution. The regularity of this series justifies greater nicety in its computation than is typical of most discrete series. Arbitrarily to give it an exact value, however, where the evidence is clear that the differences between the units arrayed (placed side by side in an ascending or descending order) are clearly unequal, is to allow the ideal position of the terms in the group to strip the median of much of its significance. This is true only if this particular form of average is considered to be more than a mathematical concept. To require that it be restricted to an actual item in an array, where the frequencies are grouped, and where exact positions are not known, is to give it a distorted but probably much more real function.[1] As a statistical instrument it seems best to consider it in the light of the material with which it is used. If, in the nature of the case, it can be located with accuracy, then so locate it; but if it can be determined only by neglecting the peculiar character of a distribution, then it is advisable to locate it only approximately.

If it is possible, by use of the median, to divide series into two equal parts, it is of course possible, by an extension of the same principle, to divide them into four or other number of equal parts. The medians dividing the halves of series into equal parts are called *quartiles*. The formula for the lower quartile — $Q1$ — *i.e.* the one below the median, is $\frac{n+1}{4}$, and for the upper quartile — $Q3$ — $\frac{3(n+1)}{4}$. A series of such measures gives a more complete picture of a distribution than can possibly be gotten from a single expression.[2]

[1] In the Dewey Report on *Employees and Wages*, the median is expressed only by group location, and this notwithstanding that the groups are small and the series exceptionally regular.

[2] More is said concerning quartiles in the chapter on *Dispersion and Skewness*.

The median is readily located graphically. In cumulative graphs or ogives it is located by bisecting the ordinate range [1] and extending a line parallel to the base until it meets the ogive and then by dropping a perpendicular at this point to the abscissa scale. Whether the median is read more accurately than by groups depends upon the considerations noted above respecting discrete and continuous series. Whether the absolute or relative frequencies are given is of no consequence. The process is the same. Moreover, the order in which the cumulating is done is immaterial. It may be on a "less than" or "more than" basis, and the data may represent a frequency or a time series.[2] In either case, it is the aggregate of the frequencies — the n — which is divided into halves. The manner in which this is done for data arranged in frequency groups is illustrated by Plate 17, by using the frequency data on pages 216–217, Chapter VII. The manner in which a time series may be divided into halves is illustrated on Plate 18, and from the following data:

[1] The variable should always be plotted on the ordinate axis.

[2] Data which admit of being cumulated from period to period, as amount of importation into a country by months or years to get a cumulated total, are illustrative.

TABLE N

TABLE SHOWING BY YEARS SINGLY AND CUMULATIVELY THE QUANTITY OF RAW COTTON IMPORTED INTO THE UNITED STATES, 1895 TO 1913, INCLUSIVE.

(*Statistical Abstract of the United States*, 1913, p. 669)

YEAR	AMOUNT OF RAW COTTON IMPORTED, IN POUNDS (*000's* omitted)		
	NON-CUMULATIVE	CUMULATIVE	
		"Up to and Including"	"After and Including"
Total . . .	1,421,152	1,421,152	1,421,152
1895	49,332	49,332	1,421,152
1896	55,350	104,682	1,371,820
1897	51,899	156,581	1,316,470
1898	52,660	209,241	1,264,571
1899	50,158	259,399	1,211,911
1900	67,398	326,797	1,161,753
1901	46,631	373,428	1,094,355
1902	98,716	472,144	1,047,724
1903	74,874	547,018	949,008
1904	48,841	595,859	874,134
1905	60,509	656,368	825,293
1906	70,964	727,332	764,784
1907	104,792	832,124	693,820
1908	71,073	903,197	589,028
1909	86,518	989,715	517,955
1910	86,037	1,075,752	431,437
1911	113,768	1,189,520	345,400
1912	109,780	1,299,300	231,632
1913	121,852	1,421,152	121,852

The first half of the raw cotton imported in the period 1895 to 1913 inclusive, came in between 1895 and approxi-

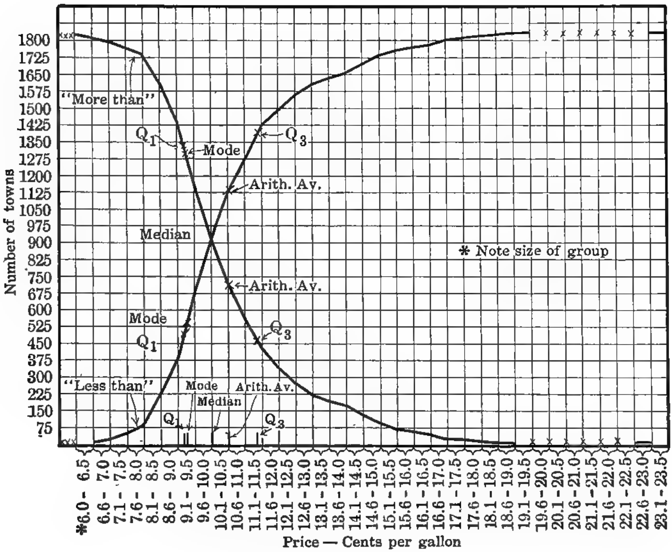


PLATE 17

Cumulative Graphs — Ogives — Constructed on "More Than" and "Less Than" Bases, Showing by Towns the Classified Prices of Oil.

mately September of 1906,[1] that is, during eleven years and eight months. The second half was imported between September, 1906, and the close of 1913, or seven years and four months. The median period — that is, the half-way period in terms of amounts imported — was September, 1906. In terms of time alone, June, 1904, is the median period. At that time, however, only 40.1 per cent of the total had been imported. These facts are shown graphically on Plate 18. In order to locate the median period in terms of importations, the ordinate axis is bisected at 710,000,000 lbs. and a line extended until it meets the historigram vertically over the period September, 1906. Obviously, in order to locate the median period in terms of time alone, the abscissa axis is bisected at June, 1904, and a perpendicular raised until it meets the historigram horizontally opposite the position 570,000,000 on the ordinate. This graphic portrayal should not be confused with that on Plate 17. In the latter case, the median *amount* is determined. In this case it is the median period or performance which is indicated. If it is desired graphically to locate the median *amount* in an historical series, amounts and not periods must be arrayed consecutively and each reported performance counted as a frequency of *one.* When this is done, the process is the same as in cumulative frequency series; that is, the amounts cumulated are plotted on the ordinate and the corresponding periods on the abscissa axis.

Objection may be raised as to the propriety of using the median for this purpose, yet there seem to be no reasons why it is not as useful and significant to divide in this manner a time as an amount or frequency concept. Indeed, in the business world, the occasion for doing the former will probably occur more frequently than the latter. Where it

[1] On the assumption of uniform importation during the year.

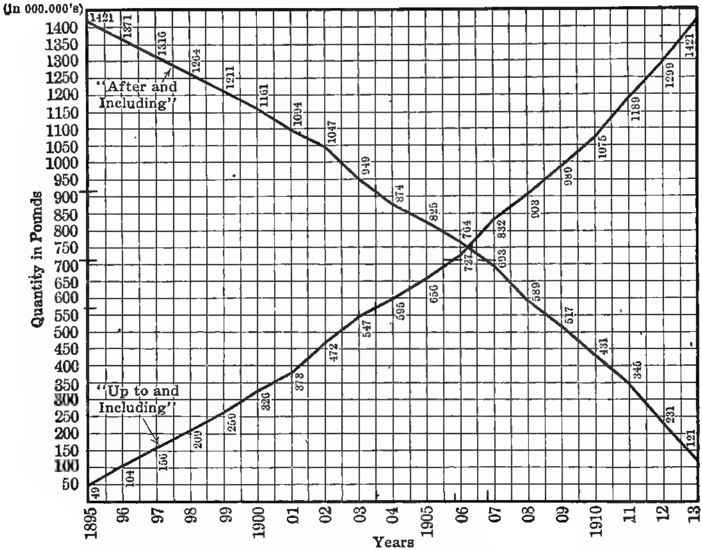


PLATE 18

Cumulative Graphs — Historigrams — Constructed on "Up to and Including" and "After and Including" Bases, Showing by Years, Importations of Raw Cotton into the United States.

is desired, for instance, to relate expenses to a definite period, the proportion attributable to one quarter or one half of the time may be of real significance. Of course, amounts, likewise, may be partitioned into equal parts and compared to the time in which incurred. In either case, by plotting the amounts cumulatively and the periods consecutively, the median *positions* may be located and related to each other.

The necessary steps in determining arithmetically the median *amount* imported are given below and the data arranged as in Table O. Place the amounts in numerical order and apply the formula $\frac{n+1}{2}$, as above. Thus, $n = 19$. $\frac{n+1}{2} = 10$, or the 10th item, which equals 70,964,000 lbs. That is, over a period of 19 years the amount imported which stood half-way between the extreme was 70,964,000 and this occurred in the year 1906. The arithmetic mean is equal to 75,800,000^{+} lbs. (The extreme items are potent here.) In this arrangement consecutiveness of amount rather than of time is followed. In the former arrangement the order is consecutive for time but not for amount.

The median as an average or summarizing expression should be used with great care. While in its computation all the frequencies are required, it is not affected by the size of the items except at or near the middle of a series. This may be a significant weakness when not only the number of times an item appears but also its positive size is important. Theoretically, it is best suited to continuous series or to discrete series in which the measurements are numerous and accurate, and when the scale is small and the groups into which they are merged narrow. It should be considered only as one measure of a complex distribution, and be compared with the arithmetic mean, and the mode whenever possible.

TABLE O

TABLE SHOWING DATA OF IMPORTATIONS OF RAW COTTON ARRANGED SO AS TO DETERMINE THE MEDIAN AMOUNT IMPORTED

PERIODS	FREQUENCIES	IMPORTATIONS IN POUNDS
Total	19	1,421,152,000
1901	1	46,631,000
1904	1	48,841,000
1895	1	49,332,000
1899	1	50,158,000
1897	1	51,899,000
1898	1	52,660,000
1896	1	55,350,000
1905	1	60,509,000
1900	1	67,398,000
1906	1	–70,964,000
1908	1	71,073,000
1903	1	74,874,000
1910	1	86,037,000
1909	1	86,518,000
1902	1	98,716,000
1907	1	104,792,000
1912	1	109,780,000
1911	1	113,768,000
1913	1	121,852,000

V. THE MODE

1. *What the Mode Is*

The mode was defined as that item in a series which is most characteristic or common. It is the typical fact, and is always represented. In the nature of the case it cannot be fictitious if its function is accurately interpreted. In series in which there is no distinct mode and where data do not

congregate, by manipulation a clearly defined one may be made to appear when in reality none exists. This is particularly true for discrete series where frequencies are widely dispersed, and where it is necessary successively to widen the groups in order to concentrate them at a particular place. The wider groups are made, however, to give a distribution regularity, the more the individuality of the data is submerged and the more unreal in discrete series does the mode become. When they are wide, it is often felt that the mode must be more accurately located than simply by group. To do this its position must be approximated by interpolation. No objection can be offered to this practice in continuous series, where measurements are merely samples and where an ideal distribution would result in case sufficient measurements were taken; but it is rarely if ever appropriate for discrete series unless measurements are numerous and tend definitely to cluster. Even when they do so, to assign the mode a definite position it is necessary to proceed arbitrarily. It should never be made to appear that there is an exact mode when there is none. The mode should be thought of as that expression which not only is a reality in itself but which really characterizes a distribution as a whole, the deviations from which shade off in a definite and regular order.

Viewed in this light, the mode has very definite limitations as an average or summarizing expression. Extreme items are entirely ignored. In this respect, it goes further than the median which assigns equal weight to all frequencies, and, of course, differs radically from the arithmetic mean. While it represents a reality, it does so only by expressing the dominant or more frequent one and ignoring the others. Moreover, there may be no mode, or there may be several modes not all of the same importance, but all sufficiently marked as to

merit attention. To ignore the lesser simply because it is lesser is never admissible. Moreover, by interpolation to make it appear that the *true* mode is located elsewhere than at the position shown is, for discrete series, inadmissible, in case the measurements are typical and sufficiently numerous. For continuous series it may be conducive to greater accuracy to widen the groups and, therefore, to remove data from the peculiarities and limitations of the units in which expressed. A distribution may be distorted by the unrepresentative character of the sampling or by the crudity of measurements.[1] The appearance of two or more modes may be due to the peculiarities of a particular set of measurements which serves only as an approximation to the real distribution for a completed series.

It must clearly be kept in mind that there are two types of distributions, the continuous and the discrete, and that the function of the mode and the ease and accuracy with which it can be located are vitally affected by this fact. Liberties which might well be taken with data of the continuous type may under no circumstances be tolerated with those which are discrete. In the former it may be legitimate to locate the mode within narrow limits, even assigning it a definite position; in the latter, except in rare cases, — and what these are is to be determined by a study of the data concerned, — the mode cannot generally be more definitely located in frequency series than by groups. In some cases, of course, discrete data tend to concentrate on definite units. When this is the case, the position of the mode is definite. If interest rates tend to concentrate on even and not on fractional per cents,[2] the modal per cents can be

[1] See data on measurements of the lengths of lobsters, Chapter V, p. 152.

[2] See Chapter V, p. 149.

located only at these places. The mode if it is anything is a reality. It is not necessarily less real, although it may appear to be, if it is sometimes assigned as a group and not as a position within a group. Indeed, the nicety with which it is located may result in making it unreal. When this is true cannot be determined by any general rule. All that can be said is that discrete and continuous series respecting the location of the mode must be viewed differently. In what way differently is determined in each case by the character of the data themselves.

2. *How the Mode is Located*

(1) The Location of the Mode in Historical Series

The thing which is modal or typical shows itself in its frequency. The exceptional is not modal. The mode is the characteristic which most frequently appears. In Table N, showing importations of raw cotton from 1895–1913, the modal year was not 1913, at which time there was imported almost three times as much cotton as there was in 1901. This is the exceptional year. Years which may be suggested as modal are 1895, 1897, 1898, 1899, 1901, and 1904, in each of which there were imported between 45 and 55 million pounds. If the conditions set up to determine the mode be altered so as to include all years in which between 45 and 60 million pounds were imported, 1896 also must be called a modal year, and 55^{+} millions a modal amount. In this case, as in so many, there is no *one* mode. The manner in which the mode may be approximated, or, more properly, perhaps, the conditions which should be imposed in its determination, may be illustrated as follows:

TABLE P

DATA SHOWING IMPORTATION OF RAW COTTON INTO THE UNITED STATES, ARRANGED SO AS TO DETERMINE THE MODAL AMOUNT

YEAR	AM'TS IN 000'S	FREQUENCIES					
		IDENTICAL COL. 1	APPROXIMATE, BY GROUPS				
			5 Mil. beginning at 45 Mil. Col. 2	10 Mil. beginning at 40 Mil. Col. 3	10 Mil. beginning at 45 Mil. Col. 4	15 Mil. beginning at 45 Mil. Col. 5	8 Mil. beginning at 46 Mil. Col. 6
1901	46,631	1					
1904	48,841	1	3	3			
1895	49,332	1			6		6
1899	50,158	1				7	
1897	51,899	1	3	4			
1898	52,660	1					
1896	55,350	1	1		2		2
1905	60,509	1	1	2			
1900	67,398	1	1				1
1906	70,964	1				5	
1908	71,073	1	3	3	4		3
1903	74,874	1					
—	—						
1910	86,037	1	2	2	2	2	2
1909	86,518	1					
—	—						
1902	98,716	1	1	1	2	2	1
1907	104,792	1	1	2			2
1912	109,780	1	1		2	2	
1911	113,768	1	1	1			1
—	—						
1913	121,852	1	1	1	1	1	1

In this table the consecutive order for amounts is followed. The grouping is: column 2, 5 million pounds; column 3,

10 million pounds; column 4, 10 million pounds, but starting at 45 million and extending to but not including 55 million; column 5, 15 million pounds; and column 6, 8 million pounds. The amounts are equally common in column 1, no account being taken of the degrees of difference in the absolute amounts. In column 2 (the grouping being 45 to 50, 50 to 55, etc.) groups 45 to 50, 50 to 55, and 70 to 75 are equally common. By widening them to 10 million pounds, as in column 3, more instances now appear at group 50–60 million than at any other place. By retaining the 10 million pound group but beginning it at 45 million, a decided concentration appears in the first group. By extending the width to 15 million, the group 45 to 60 shows the greatest concentration, but a second concentration appears in the group 60 to 75 million. Where is the mode? Undoubtedly the most characteristic amount imported when the whole period is considered is less than 60 million pounds. But how much less? The arithmetic mean of the amounts less than 60 million pounds is 50,695,000 and the median 50,158,000. The most characteristic amount with a 10 million group is 46 to 56 million, of which there are 7 instances; more narrowly, there are 5 years in which the amounts imported are between 49 and 56 million. It is probably not wise to locate the mode more accurately than in the group 46 to 54 million (column 6). To do so for this type of distribution would be to strive for too great accuracy.

For historical series, — simple historigrams, — the modal characteristic is shown graphically by the tendency for the curve to remain horizontal. Extremes in the position on the ordinate reveal exceptional conditions. By placing a ruler horizontally to the axis of abscissa and by moving it up and down, and at the same time, observing with each movement the distances covered by the graph on both the

axes, the most common characteristic of the curve and period of time over which it extends may be approximated. This is only a rough measure, but probably sufficiently accurate for historical series.

When historical data are plotted cumulatively, as in Plate 18, the modal position or positions are shown by the tendency of the graph to retain a given direction. Inasmuch as the chronological order is followed in cumulating, modal amounts may not be placed in juxtaposition and the dominant characteristic is difficult to appraise and locate. The use of the graphic method for determining the mode, so far as cumulative figures are concerned, is not advocated.

(2) The Location of the Mode in Frequency Series

When data are arranged in frequency groups, the modal position or the characteristic feature shows itself in the dominant frequency. If it is pronounced, as in Table M, the modal group may readily be distinguished. The position in the group for discrete and continuous series must be assigned in accordance with the principles discussed above. If interpolation is appropriate the position within a group is determined by giving proper weight to the frequencies on either side of it. For instance, in Table M — assuming this to be of the continuous type — 91 instances are found in the next lower group, and 49 in the next higher. Combined, they make 140 instances, $\frac{91}{140}$ of which are exerting an influence to place the mode below the group \$9.00 to \$9.99, and $\frac{49}{140}$ are exerting an influence to place it above. The actual mode is in the group \$9.00 to \$9.99. $\frac{49}{140}$ of \$1.00

— the width of the group — equals \$.35, and $\frac{91}{140}$ of \$1.00 equals \$.65. That is, the theoretical mode is \$9.00 + \$.35 = \$9.35. From the other side, the mode equals \$9.999 − \$.65 = \$9.349. If all of the frequencies on either side of the modal group are given weight, the actual mode is \$9.34 or $\frac{109}{321} + \$9.00$, or $\$9.999 - \frac{212}{321}$.

When frequency data are plotted in a simple graph, the modal position is shown by the maximum ordinate. Approach to the vertical indicates dominant frequency. The case is the reverse of that in historical graphs. Position, in respect to scale, and degree, in relation to amount, are revealed in the graphic figure. The assignment of the exact position, of course, is to be determined by the peculiarities of the data and not by the graphic portrayal of it. The latter is simply pictorial, depends upon the data, and should faithfully depict them.[1]

On ogives, or cumulative graphs, the mode or place of greatest frequency density shows where the curve passes through the greatest distance vertically and the shortest distance horizontally, *i.e.* where it is most nearly vertical. Bowley has suggested the empirical rule of rotating a ruler on the curve at this point in order to determine its exact location within the group. For most purposes the modal group is sufficiently definite for all practical purposes without this refinement. However, when a distribution approaches and recedes from the maximum very gradually, even the group position is not evident on a graph by inspection. In such cases Bowley's method may successfully be used. The positions of the modes on the distributions on Plate 17 are located in this way.

[1] Chapter VII, *passim.*

When data are arranged in frequency groups and distributions are irregular, showing no tendency to be dispersed in a definite order around a central norm, it is frequently desirable successively to widen the groups, at the same time altering the frequencies to correspond, until regularity appears. However, there is always the danger of so concealing the individual peculiarities of the data, when dealing with discrete series particularly, as to negative any real value which they may possess. Frequently, the desire for regularity of distribution is so strong that its securing is made an end. Group adjustment should properly be looked upon as a means of correcting a false impression, as for instance, when data clearly of the continuous type have been distorted, by the limitations of the units in which they are expressed or by inadequacy of sampling, from the order which they should properly assume.[1] It is always a problem to know how far to carry this synthesizing process. There is no rule-of-thumb principle which will answer the question. In effect, it is a process of smoothing and therefore, in discrete series, sacrifices individual characteristics in order to secure general impressions. The peculiarities of the whole series dominate the peculiarities of the parts. It should be remembered that for most data, particularly discrete, group widening results in a real sacrifice unless through it error is eliminated. This topic was discussed for both types of series in Chapter V, and can, therefore, be disposed of with this word of caution, and with brief reference to the following table and the corresponding graphs.

[1] See the Table showing the measurements of lengths of lobsters, Chapter V, p. 152.

TABLE Q

Table Showing the Frequency of Ratios of Building Values to Land Values for Buildings Ten Stories or More in Height, New York City, 1914

Per Cent of Building to Land Values	Frequencies by Per Cent Groups 1	2	3	4	6	8
15	2	4	5	5	8	10
16	2					
17	1	1				
18	0		3			
19	1	3		5		
20	2					
21	0	2	7		18	
22	2					
23	5	10		16		25
24	5		11			
25	5	6				
26	1					
27	1	2	6	9	15	
28	1					
29	4	7				
30	3		9			
31	2	6		9		27
32	4					
33	0	3	5		21	
34	3					
35	2	5		18		
36	3		16			
37	4	13				
38	9					
39	3	6	13	16	23	33
40	3					
41	7	10				
42	3		10			
43	0	7		17		
44	7					
45	5	10	12		24	
46	5					
47	2	5		14		30
48	3		12			
49	7	9				
50	2					
51	7	12	13	16	21	
52	5					
53	1	4				
54	3		8			
55	1	5		9		19
56	4					
57	3	4	5		14	
58	1					
59	1	5		10		
60	4		9			
61	2	5				
62	3					
63	2	2	2	3	7	12
64	0					
65	0	1				
66	1		5			
67	2	4		9		
68	2					
69	2	5	6		8	
70	3					
71	1	1		3		5
72	0		2			
73	1	2				
74	1					
75	0	1	1	2	4	
76	1					
77	0	1				
78	1		3			
79	0	2		2		2
80	2					

By successively widening the groups in which the ratios of building to land values are expressed in Table Q, it is possible to reduce the frequencies to a gradually ascending and descending order but not without destroying somewhat the peculiarities of the distribution as revealed in the column marked 1. Graphically, the result of widening the groups is shown on Plate 19.

VI. The Properties of Averages or the Average to Use [1]

Probably the properties of the different averages discussed above can more clearly be seen if the conditions are formulated which help to determine which average to use for a number of widely different cases.

Suppose we were interested in the experience of a salesman as a basis for promotion to a new territory or to an advanced wage or salary scale. The sales record of this man is given over a sufficient period, the sales being listed by territory, by grade of commodity, by prices of the article sold, by profits realized by the firm, by the length of time utilized in making them, by cost to the firm in present salary and expenses, etc., — the supposition being that the sales are in the detail that is current with the best appointed sales records. Without making an elaborate judgment on the basis of all the data listed above and such other as may be available, could one employ an average of the sales for the purpose in mind, and if so in which one could he place most reliance? Is the arithmetic mean, — an average of good and bad days, of sales among all classes of buyers, of those requiring one call and those requiring close following up, of small and large sales, of those upon which little as well as

[1] This topic is further considered in Chapter IX.

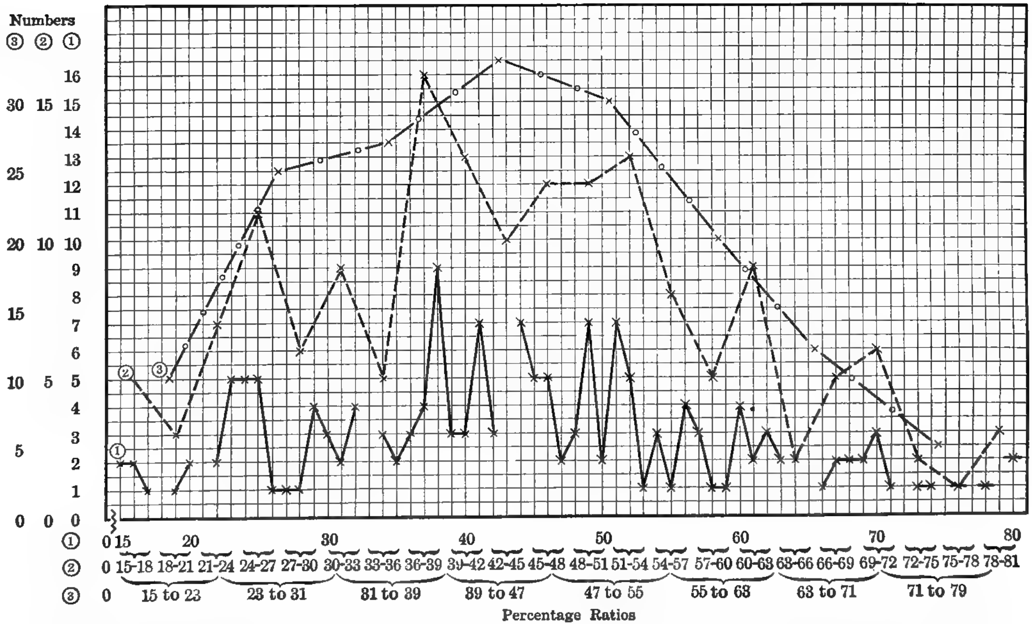


PLATE 19

Histograms Showing the Distributions of Ratios of Assessed Values of Buildings to the Assessed Values of Lands upon which they Stand, New York City, 1914.

large profits are realized, etc., to be taken as a measure of a salesman's activity, test of fitness, or worth to a company? Or are we interested in that average which takes account of the bad days and the small sales, of the good days and the large sales, but which gives no more importance to one of them than to another, realizing that the best of salesmen occasionally have off days and poor territory and that these will have to be reckoned with? Such a line of thought suggests the advisability of using the median. But, comes the retort from one who approaches the problem from another angle: "This man has had a consistent record of a high order and it is neither fair to the man nor to the company to give weight to his misfortunes. The facts show that we can expect him to make such and such a record — the overwhelming percentage of his sales are of this character; or, in other words, the percentage of the time in which he fell below a high standard is negligible and should be given no weight. If his mistakes and failures are counted, we shall be putting a premium upon mediocrity and not be giving sufficient recognition to real merit." Such an argument suggests the wisdom of using the mode as a test of fitness.

It may be argued that it is unwise to let any one set of circumstances govern, no matter from what angle the problem is approached, and, undoubtedly, this is true. However, no matter how carefully the promotion is considered, if the facts above indicated are held to be germane, it is necessary to decide upon the weight to be assigned to the approaches indicated in the various averages. It is, of course, conceivable that the various averages would not be materially different. If this is true, the case for using one at all is strengthened. As to whether averages can be used is one question: which one to use, in case they are allowable, is quite another. It is the latter question which is now

being discussed. But in this, as in many cases, a change is made, a policy is adopted irrespective of what averages show. Other aspects than the numerical are so overwhelming in importance that the case in all its bearings does not admit of statistical statement. One gets little aid from this approach.

Again, suppose that one were interested in the time necessary to reach his work, as a fact governing his location for residential purposes, and that there existed but one available means of transportation. Is it the arithmetic mean time, the median time, or the modal time in which the distance is traveled which is of interest? Delays happen even in connection with the best transportation service.[1] Should the possibility of these be considered in the allowance of time to reach one's place of employment, or should they be regarded as negligible on the ground that they are irregular and uncertain? If one sets great weight upon punctuality, he undoubtedly will allow for this factor in spite of its contingency. On the other hand, if the transportation company in question were advertising its service, it would feature the typical or modal if not the shortest performance. If the period considered were of appreciable length, it is doubtful if the differences between the various averages would be of great significance even for widely different uses. The distribution of frequencies would tend to conform to the normal law of error and the averages closely to agree. On the other hand, if the time were short and the delays at all frequent, the characteristic might be widely different from the mean time. There would be no tendency for delays to be com-

[1] See "Report" of the *Chicago Traction Subway Commission*, "On a United System of Surface, Elevated and Subway Lines," pp. 272–274, Chicago, 1916, for an analysis of the classified causes of one year's reported delays of more than five minutes' duration on the surface lines.

pensated for by exceptionally quick service, since most of the runs would be made according to scheduled time. The arithmetic mean would undoubtedly tend to be too large. It is precisely this fact which needs to be considered by the person who desires to reach his office each morning at or before a stated time, and which the advertising manager of the company desires not to bring to the attention of the public. It is evident that the averages accurately reflect the characteristics of the data, but they call attention to different things. It is this fact which is too often ignored, or at least too frequently not given sufficient attention in current discussions, in semi-scientific studies and government reports, and unfortunately in some critical studies.

One might be interested in the "average" suit of ready-made clothes turned out by a clothing concern, but the kind of an average best suited to his purposes will depend upon what those purposes are. If he is in the production side of the business his interest is in typical or standard sizes determined for him by the physical facts of size and proportion among men. The great majority of sales will be to individuals who conform within narrow limits to standard measurements. The manufacture of these garments constitutes his problem. His interest lies in the modal suit; not in the median nor in the arithmetic mean, as such. If he considered the arithmetic mean and manufactured his garments according to the sizes determined by such a calculation, it is doubtful if his customers could be fitted, since such measurements imply that the exceptionally large and the exceptionally small will affect the measurements of suits designed for the great homogeneous and standard majority. If large quantities of suits were manufactured, it is true that the mode, the median, and the arithmetic mean sizes would closely agree; but by the prudent producer this agreement

would be taken for granted only where production was on the largest scale.

Likewise, if the value instead of the size of the "average" suit were uppermost in one's mind, it is doubtful if the arithmetic mean would be particularly enlightening. Such a figure is too general, too indefinite, for any but the most superficial purposes. Some sizes tend to be normal; this grows out of a physical fact. Values tend to be normal or characteristic too, but this normality is not reflected in an arithmetic mean, as it is in the case of sizes, since all values may or may not be represented in the various sizes manufactured. Suits which can be manufactured according to set measures and in large quantities, other things being equal, tend to be cheap. Suits which are manufactured only to special order and in relatively small quantities, other things being equal, tend to be dear. The exceptional in either case would be weighted heavily and the characteristic be far different from the mean price. As a basis for roughly estimating profit an arithmetic mean price may be all that is required, but for shaping a selling policy an intimate study of the characteristic prices for the various types of demand is necessary. This is merely another way of saying that only homogeneous data can properly be averaged, and that the merits of each average must be settled in the light of its use.

The errors into which one may be led by indiscriminately using an average of non-homogeneous data are admirably shown in the following table giving deaths and death rates of married and unmarried men in Scotland.

TABLE R

Table Showing Deaths and Death Rates of Married and Unmarried Men in Scotland, 1863, Classified by Age Groups

(From the 9th Detailed Report of Dr. James Stark to the Registrar-General of Births, Deaths, and Marriages in Scotland)

Ages	Married			Unmarried		
	Number Living	Deaths	Death Rate	Number Living	Deaths	Death Rate
All ages	503,376	11,765	23.4	243,259[1]	4,189	17.2
20–25	22,946	137	6.0	106,587	1,251	11.7
25–30	54,221	469	8.7	48,618	666	13.7
30–35	66,153	600	9.1	25,962	383	14.8
35–40	63,858	690	10.8	15,857	253	16.0
40–45	62,645	782	12.5	12,311	208	16.9
45–50	54,505	869	15.9	8,824	179	20.3
50–55	49,591	880	17.7	7,636	205	26.8
55–60	38,006	929	24.4	5,550	142	25.6
60–65	35,920	1,216	33.9	5,242	227	43.3
65–70	22,021	1,134	51.5	2,848	156	54.8
70–75	16,029	1,291	80.6	2,021	205	101.4
75–80	9,716	1,135	116.8	1,081	157	145.4
80–85	5,477	953	174.0	513	101	196.9
85–90	1,708	488	285.7	151	32	211.9
90–95	449	137	305.1	50	21	420.0
95–100	103	40	388.4	6	3	500.0
100 and above	28	15	535.7	3		

[1] As reported. The correct total from the addition is 243,260. The table is quoted from Bliss, George I. — "The Influence of Marriage on the Death-rate of Men and Women," in *Quarterly Publications of the American Statistical Association,* March, 1914, p. 55.

"The first striking fact which this table reveals is that the death-rate of the bachelors was double that of the married men between the ages of 20 and 25. As its persons became older, this excessive difference in the death-rates of the married and the unmarried decreased slowly and regularly, showing the difference in favor of the married men at every period of life. It is thus proved that the state of bachelorhood is more destructive to life than the most unwholesome trades. When we come to the total death-rate at all ages, however, the very reverse is the case. The general death-rate among married men is very much higher than that among single men; so that, while only 1,723 bachelors died during the year out of every 100,000 bachelors, 2,338 married men died out of a like number of married men.

"This apparent contradiction may be explained as due to the fact that the number of bachelors being far greatest at that period of life when the mortality is very low, namely, from 20 to 24, whereas the number of married men is greatest at those periods of life when mortality is high, seeing that mortality increases with age. Furthermore, almost half of all the deaths of the bachelors occur before the thirtieth anniversary, at which period the mortality is much lower than at the more advanced periods of life. When the whole deaths at all ages are thrown together and compared with the total bachelors living, the general mortality seems to be little higher than that due to the earlier period of life. Among the married men, on the other hand, the greatest number of deaths occur between the sixtieth and seventy-fifth year of life, at which period the mortality is high as compared with the number living. Consequently, when the total deaths of husbands of all ages are compared with the total living, a high mortality seems to have prevailed, because the persons were all so much older when they died than were the bachelors. Therefore, comparing the total deaths of the married at all ages with the total deaths of the bachelors, necessarily leads to a false conclusion. In comparing mortality rates of two or more classes, to be correct, it must be limited to comparing at each age group, and the smaller we take the age group the more nearly correct are the rates." [1]

[1] *Quarterly Publications of the American Statistical Association*, March, 1914, p. 56.

While this illustration is drawn from mortality statistics, and seems to have little or no bearing on the problems of the business man, except in so far as it illustrates the error into which one may be led by making his basis of generalization too broad, and therefore his conclusion too indefinite, it suggests a problem of practical import to the business world. In most states, laws now require that employers of labor provide in some manner for the compensation of accidents which occur to their employees while engaged in the regular course of business. Through the failure to define what accidents are, and to relate those occurring to too broad a base, not differentiating between hazardous and non-hazardous occupations and between slight and severe accidents, and moreover, through the failure to keep accurate statistics of accidents, employers in this country have not had until recently, if they now have, an adequate basis for the computation of accident risk.[1] Not only have the bases been indefinite, but they have been too broad, with the result that the best that could be given was the roughest sort of a risk coefficient — a crude average without practical merit. Discrimination as between severe and minor accidents, and hazardous and non-hazardous conditions of employment, is the first essential to clear thinking about accidents, and the first guaranty of the reasonableness of insurance premiums.[2] A rough arithmetic mean, a median, or a mode, per se, is not enough. What is necessary is the determination of the characteristic accident rate, not for industries as a group, but for conditions of employment, definitely standardized, within each industry.

Statistics should always relate to definite conditions and

[1] Rubinow, I. M., "The Standard Accident Table as a Basis for Compensation Rates," *Quarterly Publications of the American Statistical Association,* March, 1915, pp. 358–415.

[2] *Ibid.*, pp. 358 ff.

circumstances. Duplicate these and the statistical facts are likely to be repeated. Alter them and the consequences are different. Before a policy can be mapped out on the basis of statistical facts alone, or given consequences said to follow from given conditions, the latter must definitely and clearly be defined and their boundaries indicated. So-called statistical laws operate with implacable regularity only when conditions producing them occur with unchanging persistence. To establish beyond cavil cause and effect requires not only that statistical data be referred solely to the conditions that produce them, but also that the statistical means employed to interpret them be appropriate to the purposes in mind. To assign meaning to averages alone without taking the trouble to determine the conditions which produce them or their suitability to the cases in point is as wrong statistically as to draw a false analogy logically. To do the first is to ignore the existence of determining circumstances; to do the latter to ignore their application.

"An average is not to be regarded as a secret something which determines events. This blunder is often made in social statistics. After finding a certain average in human affairs, we conclude that some secret fate is at work. By the aid of a little rhetoric we easily persuade ourselves that an event is fully accounted for when 'the law of averages' demands it. 'There may be an average in birth and death and crime, but, after all, the average is not responsible for any of them. It takes something more potent than an average to produce typhoid fever or to crack a safe.'" [1]

To employ an average suggests the formation of a judgment or a conclusion following from a full consideration of detail which it replaces. An average represents the culmination of a process of thought which when removed from the steps required for its determination is likely to be assigned

[1] Coffey, P., *The Science of Logic*, Vol. II, p. 291.

new meanings and used for purposes foreign to those for which it was designed. Given statistical application, this means that chronologically averages come late in the process of analysis. They should be used with discrimination as to function and in close contact with supporting detail, with the realization that they emphasize the generalizations and comparisons which seem to be warranted after a careful and painstaking scrutiny of the problem from the angle from which it is approached.[1]

The functions of averages are unmistakable; the justification of employing them must be determined by an appeal to all the facts and in the light of the peculiarities characteristic of the different types. As a statistical caution let it be said: *Do not rush headlong into the uses of averages. They are commonly but vaguely understood, and it is the particular function of the statistician to adopt that caution and circumspection in the use of numerical facts which the seeming exactness of his tools appear not only to suggest but to make imperative.*

[1] "But however often an average may have been confirmed, we can never attribute to it the importance of being by itself the expression of any necessity. Every result is necessary when its conditions are given; every particular instance was necessary in so far as from the given conditions it could only be such and no other; all individual determinations and differences in the particular cases, which were neglected by the average, were necessary; the most extreme deviations were necessary, and it will also be necessary, if all the particular conditions recur in exactly the same way, that they should again have the same results, and that therefore the sum of the results will be the same. . . .

"Such uniformities of numbers and averages are primarily mere descriptions of facts which need explanation as much as the uniformity of the alteration between day and night; and the explanation can be found only where the actual conditions . . . are forthcoming. But these are the concrete conditions of the particular instances counted, they are not directly causes of the numbers; it is only the nature of the concrete causes which can show it to be necessary for the effects to appear in certain numbers and numerical relations." Sigwart, C., *Logic*, Vol. II, p. 490.

VII. Summary and Conclusion

An average should be considered as derivative and as summarizing and characterizing data in a single expression.[1] The average best suited for a particular use depends upon the purpose one has in mind. Frequently, it is desirable and necessary to compute not only the arithmetic mean but the median and mode in order to safeguard oneself against criticism and to reflect types of distributions more in detail. The relative stability which these averages assume is enlightening in itself. If it is remembered that the computation of the arithmetic mean and the median requires all the frequencies; that the former is affected by both the size of items and frequencies, while the latter is affected by frequencies and not by the size of items except those at or near the middle; and further, that in the computation of the mode both the size and frequencies of exceptional items are ignored, it is evident that in changing the order or number of frequencies the mode is scarcely affected at all; that the median is only slightly affected, and the arithmetic mean violently affected.

No single average will suffice for all purposes. Each is affected differently by arrangement, frequency, and size of items, and should be used with a full knowledge of the peculiarities of distributions. One is never justified in employing a short-cut expression in order to describe a complex

[1] An average "is an abbreviation, and it has so much in common with the ordinary logical abstract concept that it neglects all differences, and we cannot tell from it how far the numbers from which it is obtained, or which it has to represent, may differ from each other. It is, however, inferior to the general concept in so far as the latter is a statement of what is the same in all the particular instances, while the average is merely a fictitious value which may never actually occur in any particular case, and which by itself does not even justify us in expecting that the majority of the particular instances in a region will approximate to it." Sigwart, C., *Logic*, Vol. II, p. 487.

whole unless he realizes the limitations of the instrument which he uses. Too frequently averages are used or computed without realizing their limitations and appreciating the fact that there is a best average to employ. Derivative expressions of this character are often imperfect substitutes for detail. Frequently, an exceptional instance which would be ignored in the use of the mode is that particular instance in which one has greatest interest. On the other hand, the inclusion of an exceptional item in determining the arithmetic mean may serve to so prejudice it as to give a wholly erroneous picture of the characteristics which are dominant. The average to be used is invariably a function of the purpose which one has in mind. If that purpose is to complete a vivid and well-rounded picture of a complex thing, a single stroke, as it were, in the form of the average, notwithstanding the fact that it is included within the picture, will not suffice when a vivid and concise description is necessary. As classified data are more readily understood and compared than those in heterogeneous form, and tabular arrangement superior to unscientific classification, so summary expressions of complex situations in the form of averages are frequently more significant than the detail. The passage, however, from the particular to the general — that is, from details to averages — offers precisely the opportunity for eliminating the peculiar and significant features of discrete series. In the case of continuous series the conditions are somewhat different. As the widening of groups may result in a more accurate expression of a general tendency or an ideal distribution, so a more accurate expression of a complex whole may result from the use of a single unit, as mean, median, or mode.

Caution, foresight, and analysis are necessary at every step in the use of averages — caution as to the averages to

be employed, foresight as to the meaning which may be attached to them, and analysis as to the possibilities of data to be characterized in such a manner. The following tests should always be applied. Is it possible to employ a single expression to depict the details which are essential in order to view the data in all their bearings? Is the greatest interest in the characteristic feature, in the median position, or in that mathematical position at which the arithmetic mean falls? Is it necessary to employ all these descriptive units? No single answer to these various inquiries can be given. The use of an average may be legitimate and still the question as to the most appropriate average be left in doubt. One cannot answer the first question, as it were, by intuition. Data must be analyzed and the functions of averages in general and in particular clearly be perceived before answer can be given. As caution and analysis are necessary in the employment of averages, so discrimination and judgment are necessary in assigning importance to them when used by others.

A fitting close to the discussion of averages is found in the words of Dr. John Venn. "Every sort of average — and there are many such sorts — is a single fictitious substitute of our own for the plurality of actual values existent in the results which are naturally or artificially set before us. It is impossible, therefore, for the former, in any case, effectually to take the place of the latter. But the extent to which it may succeed or fail in doing so will depend upon the nature of the facts presented to us, and still more upon the precise object we have in view." [1]

[1] Venn, Dr. John, "On the Nature and Use of Averages," *Journal of the Royal Statistical Society*, Vol. IV, 1891, p. 447.

References

Bowley, A. L. — *Elements of Statistics*, Ch. V, pp. 107–130; Ch. VI, pp. 131–143.

——— *An Elementary Manual of Statistics*, Chs. III and IV, pp. 15–35 (especially Ch. III).

Elderton, W. P. and E. M. — *Primer of Statistics*, Chs. I and II, pp. 1–23.

King, W. I. — *Elements of Statistical Method*, Ch. XII, pp. 121–140.

Secrist, H. — "The Use of Averages in Expressing the Wages and Hours of Milwaukee Street Car Trainmen," *Publication American Statistical Association*, Vol. 13, pp. 279–298 (1912).

Yule, G. Udney — *Introduction to the Theory of Statistics*, Ch. VII, pp. 106–132.

Žižek, Franz — *Statistical Averages*, Part I, "Nature and Purposes of Averages," Ch. VI, pp. 92–127; Part II, "The Arithmetic Mean," Ch. II, pp. 138–163; Part II, "The Median," Ch. IV, pp. 199–221; Part II, "The Mode," Ch. V, pp. 222–247.

CHAPTER IX

THE PRINCIPLES OF INDEX NUMBER MAKING AND USING

I. Introduction

Both business men and students of economics have come to look upon index numbers as ever ready tools for measuring, among other things, price and wage changes. In their search for summary figures for comparing distant times and far-removed places in the interest of employers, employees, or the public as consumer, producer, or investor, recourse is had to regularly prepared index numbers. They seem to possess that generality of application and representativeness of conditions that are demanded by those who are ever ready to make bold comparisons and to draw sweeping conclusions. Only rarely is time given to a consideration of the source of data, to the methods by which an index number is computed, and to the problem of how fully in a narrow sense it really serves the purpose that it is made to fit. Its composite character ought generally to be sufficient to suggest caution and consideration. The fact that it is concerned with such elusive and indeterminate things as prices of commodities and services ought to be sufficient warning against hasty use. However, in the hands of the business man, this fact often serves only to give him more confidence inasmuch as the conditions governing the prices of things in which he deals, he seems to know so well. This, however, is far from an adequate guaranty against improper use of, or positive proof that

his knowledge applies to the prices used in this case. The specially prepared index number too frequently becomes in his hands a "general purpose" index number — or probably the reverse is as often true — and, of course, as such, is open to all the limitations characteristic of general and indefinite summarizing expressions or of loose and ill-defined terms. It is to warn against such practice as well as to develop the principles of index number making that two chapters are devoted to index numbers. This chapter is concerned with a discussion of the principles involved in their construction and use; the following one to a description and comparison of the more common American index numbers.

II. What Index Numbers Are

Index numbers may for the present be defined as relative numbers in which data for one year or other period, or an average for a year or other period, are taken as a base, generally indicated as 100, and upon which data for subsequent years or other periods are computed as percentages. Until recently such numbers have been almost exclusively averages of relatives. That is, in the case of a price index, prices for subsequent periods have been expressed as relatives of prices in a base period, and averages of these taken as index numbers for the various periods. Now, however, several reputable index numbers are being computed as the sum of actual prices. In many respects this change seems desirable. These topics are discussed below in detail.

The method of computing a simple average of relative prices index number is illustrated in the following table. The first part gives the average wholesale prices of certain commodities as reported by the United States Bureau of Labor Statistics for the years 1912, 1913, and 1914. The

second part contains the same prices reduced to a relative basis, using 1912 prices as a base, as well as the averages which are the index numbers for the various years.

TABLE A

TABLE GIVING DATA FOR THE COMPUTATION OF A SIMPLE AVERAGE OF RELATIVE PRICES INDEX NUMBER

ABSOLUTE

COMMODITIES	AVERAGE PRICES IN		
	1912	1913	1914
Corn, cash, contract grades, per bu.	$ .6855	$ .6251	$ .6953
Cotton, Upland Middling, New York, per lb.	.1150	.1279	.1210
Oats, cash, per bu.	.4380	.3758	.4191
Hay, Timothy No. 1, per ton	20.4104	16.0288	15.6863
Hides, green salted, packers'; heavy native steers, per lb.	.1760	.1839	.1963
Cattle, steers, choice to prime, per 100 lbs.	9.3585	8.9288	9.6520
Hogs, heavy, per 100 lbs.	7.5954	8.3654	8.3608

RELATIVE

Total of relatives	700	676.7	704.0
Index Numbers or Averages of Relatives	100	96.7	100.6
Corn (as above)	100	91.2	101.5
Cotton (as above)	100	111.2	105.2
Oats (as above)	100	85.8	95.7
Hay (as above)	100	78.5	76.8
Hides (as above)	100	104.5	111.5
Cattle (as above)	100	95.4	103.2
Hogs (as above)	100	110.1	110.1

While index numbers have been largely restricted to price phenomena, this is by no means necessary. Any phenomenon extending over a period of time and expressed numerically may be put in this form, the only peculiarity being that its relative rather than its absolute aspect is exhibited. Index numbers of wages, of rents, of imports or exports, sales, or of any other phenomena may be constructed. Historically, price indexes were the first to be computed and to these our major attention is given, inasmuch as they are currently compiled and are those in which the business man and student of economics probably have most interest.

The purpose of an index number is to reduce to a common denominator the qualities of different factors or phenomena so as to allow comparison generally historically. It is to translate absolute into relative qualities in order that comparisons may be made. Moreover, index numbers are summaries direct or indirect of things having a common quality, as for instance, in the case of price indexes, a selling value. They represent this quality as an aggregate or average at different times for purpose of comparison. If they are aggregates of prices rather than averages of relative prices, they are no less averages. They represent divergent things, responding differently to conditions of price determination and occupying different positions in the economy of business. Being aggregates or averages, they do not in themselves reveal all the peculiarities of the things which go to make them up.[1] If averages may be fictitious and unreal,

[1] " . . . it must be borne in mind that no index number corresponds to a real thing. It is not like the mean of certain observations in natural science — such, for example, as those for measuring the distance between the earth and the sun — of which any one may err, but whose average will point to a single specific fact. An index number points to no single fact. It gives, to repeat, only an indication of a general trend of prices. People often think and speak loosely on this topic, as if an index number told the whole story once for all. There is no one change in prices. There is a

giving no evidence of the characteristic features of their several parts, and it becomes necessary to study the parts in order to understand the aggregate; or, on the other hand, if they may be real in every sense of the word, inasmuch as they represent the mode or characteristic thing, without at the same time revealing it, — so may index numbers be fictitious and unrepresentative for one use but well suited for another. Everything depends upon the purpose for which they are computed and the factors which are important in their make-up. Blindly to employ a consumer's index number in a problem relating to capital investment is a practice of the same sort as to use an average in blind indifference to the things which go to make it up. The same is true respecting index numbers of wages, of rents, or of any other thing. Realizing the importance of this truth and in consistency with what has gone before, a large part of this chapter is devoted to the principles of index number making.

III. The Uses and Computation of Index Numbers

In what has gone before emphasis has been put on plan and purpose in statistical study. These need to be insisted upon especially in connection with this topic, because, while most index numbers are of the "general purpose" type, they are given particular use.

"Few of the widely-used index numbers, . . . are made to serve one special purpose. On the contrary, most of them are 'general-purpose' series, designed with no aim more definite than that of measuring changes in the price level. Once published they are used for many ends — to show the depreciation of gold, the rise in the cost of living, the alternations of business prosperity and de-

medley of many changes, different in direction and degree. All that we can hope to secure by averaging and summarizing is some concise statement of the general drift." Taussig, F. W., *Principles of Economics*, Vol. I, p. 294. (Revised Edition, 1915.) Macmillan, New York.

pression, and the allowance to be made for changed prices in comparing estimates of national wealth or private income at different times. They are cited to prove that wages ought to be advanced or kept stable; that railway rates ought to be raised or lowered; that 'trusts' have manipulated the prices of their products to the benefit or the injury of the public; that tariff changes have helped or harmed producers or consumers; that immigration ought to be encouraged or restricted; that the monetary system ought to be reformed; that natural resources are being depleted or that the national dividend is growing. They are called in to explain why bonds have fallen in price and why interest rates have risen, why public expenditures have increased, why social unrest prevails in certain years, why farmers are prosperous or the reverse, why unemployment fluctuates, why gold is being imported or exported, and why political 'landslides' come when they do."[1]

Generally, however, two major purposes are distinguishable, so far as price indexes are concerned. First, that of measuring quantitatively change in price level from time to time, and second, that of interpreting the effect of change upon various types of people. The first index number (or use) is often called the Jevonian, because the English economist Jevons was among the first to attempt to measure the change in the purchasing power of gold. The second index number (or use) — hardly a type of index number, although the conditions of its computation are somewhat different from those which characterize the first — is the so-called *consumers'*. Its purpose is to approximate the effect of price changes upon consumers. Of course, there might, with the same justice, be computed a "producers'" index number, the only difference being that emphasis would be placed on other commodities — those in which they are interested and which enter into their costs.

[1] Mitchell, Wesley C., "Index Numbers of Wholesale Prices in the United States and Foreign Countries," *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173, July, 1915, pp. 25–26.

Inasmuch as few students or business men have the necessary time and organization to construct index numbers suited to their particular purposes, and because there are now currently published many price index numbers, the order in which our discussion has proceeded — that is, from definition of purpose to employment of data — is reversed. The purpose for which index numbers may be used must be settled in the light of the peculiarities of the numbers at hand. This calls for detailed and intimate study, and must follow the lines suggested in this chapter.[1]

Professor Mitchell enumerates the operations involved in making a price number as follows:

> "(1) Defining the purpose for which the final results are to be used; (2) deciding the numbers and kinds of commodities to be included; (3) determining whether these commodities shall be treated alike or whether they shall be weighted according to their relative importance; (4) collecting the actual prices of the commodities chosen, and, in case a weighted series is to be made, collecting also data regarding their relative importance; (5) deciding whether to measure the average variations of prices or the variations of a sum of actual prices; (6) in case average variations are to be measured, choosing the base upon which relative prices shall be computed; and (7) settling upon the form of average to be struck.
>
> "At each one of these successive steps choice must be made among alternatives that range in number from two to thousands. The possible combinations among the alternatives chosen are indefinitely numerous. Hence there is no assignable limit to the possible varieties of index numbers, and in practice no two of the known series are exactly alike in construction. To canvass even the important variations of method actually in use is not a simple task."[2]

[1] Such a comparative study has been made by Professor Wesley C. Mitchell in "Index Numbers of Wholesale Prices in the United States and Foreign Countries," *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173, July, 1915. Acknowledgments are here made of the indebtedness of the writer to Professor Mitchell for much of the illustrative matter in this and the following chapter.

[2] *Ibid.*, p. 25.

1. *Data from which Price Index Numbers are Made*

In a study of prices attention must first be centered upon the commodities included and the conditions of price making. Distinction will have to be made between producers' and consumers' goods,[1] between raw and manufactured commodities,[2] between manufactured goods bought by consumers for

[1] " . . there are characteristic differences between the price fluctuations of manufactured commodities bought by consumers for family use and the price fluctuations of manufactured commodities bought by business men for industrial or commercial use. . . . Though consisting more largely of the erratically fluctuating farm products, the consumers' goods are steadier in price than the producers' goods, because the demand for them is less influenced by changes in business conditions." *Op. cit.*, pp. 60–61.

[2] "These several comparisons establish the conclusion that manufactured goods are steadier in price than raw materials. The manufactured goods fell less in 1890–1896, rose less in 1896–1907, again fell less in 1907–1908, and rose less in 1908–1913. Further, the manufactured goods had the narrower extreme range of fluctuations, the smaller average change from year to year, and the slighter advance in price from one decade to the next. It follows that index numbers made from the prices of raw materials, or of raw materials and slightly manufactured products, must be expected to show wider oscillations than index numbers including a liberal representation of finished commodities." *Op. cit.*, p. 53.

"First, the list of commodities used by the Bureau of Labor Statistics includes 29 quotations for iron and its products, 30 quotations for cotton and its products, and 18 for wool and its products, besides 8 more quotations for fabrics made of wool and cotton together. On the other hand it has but 7 series for wheat and its products, 8 for coal and its products, 3 for copper and its products, etc. The iron, cotton, and wool groups together make up 85 series out of 242, or 35 per cent of the whole number. . . . Similarly, cotton, wool, and wheat, or coal, or cattle, with their products, make 20 per cent of the series in the third index number.

"Does this large representation of three staples distort these index numbers — particularly the bureau's series where the disproportion is greatest? Perhaps; but if so the distortion does not arise chiefly from the undue influence assigned to the price fluctuations of raw cotton, raw wool, and pig iron. For, contrary to the prevailing impression, the similarity between the price fluctuations of finished products and their raw materials is less than the similarity between the price fluctuations of finished products made from different materials. . . . As babies from different families are more like one another than they are like their respective parents, so here the relative prices of cotton textiles, woolen textiles, steel tools, bread, and shoes differ far less among themselves than they differ severally from the relative prices of raw cotton, raw wool, pig iron, wheat, and hides. Hence the inclusion of a large number of articles made from iron, cotton, and wool

family use and manufactured commodities bought by business men for industrial uses,[1] between mineral products, animal products and farm crops,[2] etc., the prices of all of which respond differently to conditions of scarcity and surplus.[3]

affects an index number mainly by increasing the representation allotted to manufactured goods. What materials those manufactured goods are made from makes less difference in the index number than the fact that they are manufactured. To replace iron, cotton, and woolen products by copper, linen, and rubber products would change the results somewhat, but a much greater change would come from replacing the manufactured forms of iron, cotton, and wool by new varieties of their raw forms." *Op. cit.*, pp. 61–63.

[1] "It has been found that among manufactured commodities those bought for family consumption are steadier in price than those bought for business use." *Op. cit.*, p. 64.

[2] "Third, there are characteristic differences among the price fluctuations of the groups consisting of mineral products, forest products, animal products, and farm crops. . . . Fifty-seven commodities are included, all of them raw materials or slightly manufactured products. Here the striking feature is the capricious behavior of the prices of farm crops under the influence of good and bad harvests. The sudden upward jump in their prices in 1891, despite the depressed condition of business, their advance in the dull year 1904, their fall in the year of revival 1905, their failure to advance in the midst of the prosperity of 1906, their trifling decline during the great depression of 1908, and their sharp rise in the face of reaction in 1911 are all opposed to the general trend of other prices. The prices of animal products are distinctly less affected by weather than the prices of vegetable crops, but even they behave queerly at times, for example in 1893. Forest-product prices are notable chiefly for maintaining a much higher level of fluctuation in 1902–1913 than any of the other groups, a level on which their fluctuations, when computed as percentages of the much lower prices of 1890–1899, appear extremely violent. Finally, the prices of minerals accord better with alternations of prosperity, crisis, and depression than any of the other groups. And the anomalies that do appear — the slight rise in three years (1896, 1903, and 1913) when the tide of business was receding — would be removed if the figures were compiled by months. For the trend of mineral prices was downward in these years, but the fall was not so rapid as the rise had been in the preceding years, so that the annual averages were left somewhat higher than before. An index number composed largely of quotations for annual crops, then, would be expected at irregular intervals to contradict capriciously the evidence of index numbers in which most of the articles were mineral, forest, or even animal products." *Op. cit.*, pp. 53 and 58.

[3] This topic has been given elaborate treatment by Professor Mitchell in his *Business Cycles* (University of California, Memoirs, Vol. III, September, 1913), pp. 93–109.

Obviously, a price index number which reflects price changes at large must be made from samples of all commodity groups that are affected in a peculiar manner. On the contrary, in using an index number prepared by another, one must satisfy himself respecting the list of commodities used before he can be sure what in reality the index measures.

But what is meant by "price"? Has one in mind retail or wholesale price? price at what place? under what condition of sale? to whom? price of what grade of commodity? on what market? Are price data extant? will they continue to be available? Are the "prices" contract, import, or market prices? What is *the* wholesale or retail price of a commodity?

"We commonly speak of *the* wholesale price of articles like pig iron, cotton, or beef as if there were only one unambiguous price for any one thing on a given day, however this price may vary from one day to another. In fact there are many different prices for every great staple on every day it is dealt in, and most of these differences are of the sort that tend to maintain themselves even when markets are highly organized and competition is keen. Of course varying grades command varying prices, and so as a rule do large lots and small lots; for the same grade in the same quantities, different prices are paid by the manufacturer, jobber, and local buyer; in different localities the prices paid by these various dealers are not the same; even in the same locality different dealers of the same class do not all pay the same price to every one from whom they buy the same grade in the same quantity on the same day. To find what really was the price of cotton, for example, on February 1, 1915, would require an elaborate investigation, and would result in showing a multitude of different prices covering a considerable range.

"Now the field worker collecting data for an index number must select from among all these different prices for each of his commodities the one or the few series of quotations that make the most representative sample of the whole. He must find the most reliable source of information, the most representative market, the most

typical brands or grades, and the class of dealers who stand in the most influential position. He must have sufficient technical knowledge to be sure that his quotations are for uniform qualities, or to make the necessary adjustments if changes in quality have occurred in the markets and require recognition in the statistical office. He must be able to recognize anything suspicious in the data offered him and to get at the facts. He must know how commodities are made and must seek comparable information concerning the prices of raw materials and their manufactured products, concerning articles that are substituted for one another, used in connection with one another, or turned out as joint products of the same process. He must guard against the pitfalls of cash discounts, premiums, rebates, deferred payments, and allowances of all sorts. And he must know whether his quotations for different articles are all on the same basis, or whether concealed factors must be allowed for in comparing the prices of different articles on a given date."[1]

If it is difficult to establish *the* price of a commodity at *one* time it is even more difficult to guarantee that *the* price determined at one time is *the* price at some *other* time. Conditions of marketing change, commodities change as to quality and salability, and price lists of identical commodities for any great length of time are frequently not available. The paucity of price data and the unwillingness of people to place any reliance in those extant were undoubtedly the main reasons for the relatively late development of index numbers.[2]

To-day, of course, such data as those from which the index number currently published by the United States Bureau of Labor Statistics is computed, are furnished by reputable firms and corporations, according to uniform instructions, on uniform blanks, and are carefully scrutinized by the agents of the Government. Even under these circumstances, the Bureau found it necessary to resort to a questionable statistical method of conversion in order to maintain the identity of the

[1] *Op. cit.*, pp. 27–28. [2] *Op. cit.*, p. 9.

index number, and finally radically to readjust its method of computation so as to admit new commodities in the place of those which ceased to be quoted or which became of less importance than others which ought to have been included.

But how many commodities are necessary in order that an index number may indicate either the amount or effect of price change? From what regions should prices be drawn, and how frequently ought they to be recorded? Are prices quoted in standard and definite units?[1] Some commodities are sensitive to conditions of demand and supply; others react slowly under changed conditions. Some are vitally affected by seasons, while others show appreciable change only in the face of violent disturbance and exhibit a steady rise or fall only over long periods. "Typical" price behavior can hardly be predicted for any commodity. It may never occur.

What principles have been followed in the choice of commodities? Are raw and manufactured commodities disproportioned? Is a certain unimportant commodity for one purpose — or important for another — represented in both its raw and its manufactured state? How is the importance of a commodity given weight? What test of importance is applied? how is it measured? These are vital questions which one must answer for himself for every index number before he uses it for a particular purpose.[2]

[1] "Often the form of quotation makes all the difference between a substantially uniform and highly variable commodity. For example, prices of cattle and hogs are more significant than prices of horses and mules, because the prices of cattle and hogs are quoted per pound, while the prices of horses and mules are quoted per head." *Op. cit.*, p. 45.

[2] Both for American and European index numbers such questions as these and many more are answered in *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173, to which reference has so frequently been made.

"Difficult as it is to secure satisfactory price quotations, it is still more difficult to secure satisfactory statistics concerning the relative importance of the various commodities quoted. What is wanted is an accurate census of the quantities of the important staples, at least, that are annually produced, exchanged, or consumed. To take such a census is altogether beyond the power of the private investigators or even of the Government bureaus now engaged in making index numbers. Hence the compilers are forced to confine themselves for the most part to extracting such information as they can from statistics already gathered by other hands and for other purposes than theirs. In the United States, for example, estimates of production, consumption, or exchange come from most miscellaneous sources: From the Department of Agriculture, the Census Office, the Treasury Department, the Bureau of Mines, the Geological Survey, the Internal Revenue Office, the Mint, associations of manufacturers or dealers, trade papers, produce exchanges, traffic records of canals and railways, etc. The man who assembles and compares estimates made by these various organizations finds among them many glaring discrepancies for which it is difficult to account. Such conflict of evidence when two or more independent estimates of the same quantity are available throws doubt also upon the seemingly plausible figures coming from a single source for other articles. To extract acceptable results from this mass of heterogeneous data requires intimate familiarity with the statistical methods by which they were made, endless patience, and critical judgment of a high order, not to speak of tactful diplomacy in dealing with the authorities whose figures are questioned."[1]

Mitchell, following an elaborate comparison of the various American index numbers, so far as choice of commodities and the importance assigned them are concerned, arrives at the following conclusions:

"As for the small series made from the prices of foods alone or from the prices of any single group of commodities, it is clear that, however good for special uses they may be, they are untrustworthy as general-purpose index numbers."[2]

[1] *Op. cit.*, p. 28. [2] *Op. cit.*, p. 68.

"The second conclusion . . . is that large index numbers are more trustworthy for general purposes than small ones, not only in so far as they include more groups of related prices, but also in so far as they contain more numerous samples from each group. What is characteristic in the behavior of the prices of farm crops, of mineral products, of manufactured wares, of consumers' goods, etc. — what is characteristic in the behavior of any group of prices — is more likely to be brought out and to exercise its due effects upon the final results when the group is represented by 10 or 20 sets of quotations than when it is represented by only one or two sets. The basis of this contention is simple: In every group that has been studied there are certain commodities whose prices seldom behave in the typical way, and no commodities whose prices can be trusted always to behave typically. Consequently, no care to include commodities belonging to all the important groups can guarantee accurate results, unless care is also taken to get numerous representatives of each group."[1]

2. *Methods of Computing Price Index Numbers*

In the discussion of the choice of commodities and of the difficulties of getting adequate prices the question of the method of computation has not been raised. Tentatively in defining index numbers, however, they were spoken of as relative numbers calculated upon a base, and most generally as averages of relatives. We have now to discuss the questions of the base, the amount of weight which is assigned to various types of commodities, and whether an average of relatives seems to possess any merits over the more simple aggregate of prices. Before doing so, however, some attention should be given to the peculiarities of price fluctuations.[2]

[1] *Op. cit.*, pp. 70–71.

[2] In this discussion a price index is used for purposes of illustration. The treatment follows very closely that of Wesley C. Mitchell in *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173.

(1) Peculiarities of Price Fluctuations

The trend of price change is generally in one direction for a considerable period. There are periods of falling and of rising prices. This, of course, does not mean that all prices change in the same direction at the same time, nor that those which change together change in the same degree. All that is meant is that in terms of a single year or an average of years taken as a base the price level moves up or down through relatively long periods. The differences of prices from the norm, whether negative or positive, generally tend to be in the same direction. Large differences, of course, are less common than small ones, but those that are positive do not exactly compensate for those that are negative. Mitchell has shown this in a striking way by comparing the price variations of 241 commodities in 1913, computed, first, as percentages of rise or fall from the prices in 1912; and second, as percentages of rise or fall from the average prices of 1890–1899. Graphically, Plate 20 [1] reveals the result.

The differences — excesses and deficiencies of the percentages of the 1913 prices in terms of the 1912 prices — arrange themselves, as shown by the solid line, about a norm, the arithmetic mean, the mode and the median tending closely to agree.

"But the distribution of the second set of variations (percentages of change from the average prices of 1890–1899) as represented by the area inclosed within the dotted line belongs to a different type. It has no pronounced central tendency; it shows no high degree of concentration around the arithmetic mean (+ 30.4 per cent) or median (+ 26 per cent). It is more like an oblong than like the bell-shaped normal curve; it has a range between

[1] *Op. cit.*, p. 22.

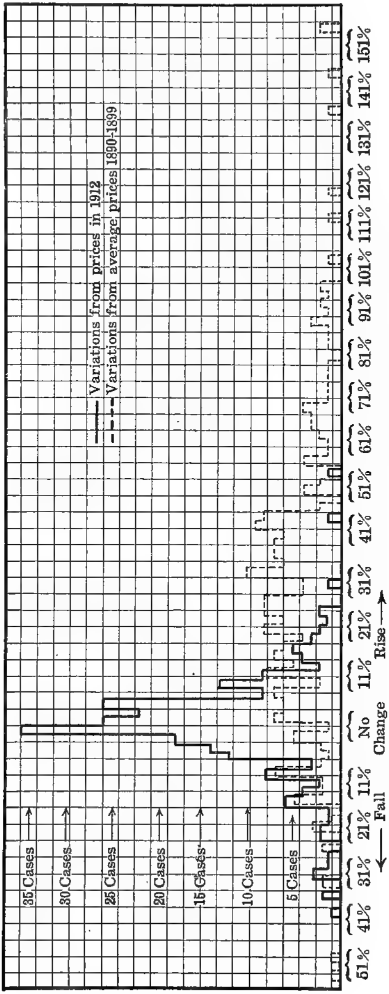


PLATE 20

Distribution of the Price Variations of 241 Commodities in 1913.
(Percentages of Rise or Fall in Prices)

the greatest fall (52.2 per cent) and greatest rise (234.5 per cent) so extreme that two of the cases could not be represented on the chart; and its probable deviation is five times as great as that of the corresponding variations from 1912 prices — 18.5 points as against 3.6.

"Price variations, then, become dispersed over a wider range and less concentrated about their mean as the time covered by the variations increases. The cause is simple: With some commodities the trend of successive price changes continues distinctly upward for years at a time; with other commodities there is a consistent downward trend; with still others no definite long-period trend appears. In any large collection of price quotations covering many years each of these types, in moderate and extreme form, and all sorts of crossings among them, are likely to occur. As the years pass by the commodities that have a consistent trend gradually climb far above or subside far below their earlier levels, while the other commodities are scattered between these extremes. Thus the percentages of variation for any given year gradually get strung out in a long, thin, and irregular line, without a marked degree of concentration about any single point." [1]

The tendency for price changes calculated from year to year, to arrange themselves around a central tendency — to conform to the "normal law of error" — has been worked out by Mitchell for the years 1891–1913, for 5578 cases. That is, the prices for more than 230 commodities during this period were expressed as percentages of the price which each bore in the preceding year, thus giving a detailed account of how each operated each year in terms of the preceding year. The changes were arranged in ascending order from the greatest decrease up through no change to the greatest increase. For the extreme distribution decils were then worked out for each year. A study of the data makes it possible to measure the concentration about a norm and to indicate the differences

[1] *Op. cit.*, p. 23.

by successive decils. Mitchell's table revealing this fact is given in the note below.[1]

[1] AVERAGE CONCENTRATION OF PRICE FLUCTUATIONS AROUND THE MEDIAN, 1891 TO 1913

[The fluctuations represent percentage changes from average prices in the preceding year.]

AVERAGE RANGE COVERED BY THE —										
1st and 10th tenths of the price fluctuations	2d and 9th tenths of the price fluctuations	3d and 8th tenths of the price fluctuations	4th and 7th tenths of the price fluctuations	5th and 6th tenths of the price fluctuations	Successive tenths of the price fluctuations	Central two tenths of the price fluctuations	Central four tenths of the price fluctuations	Central six tenths of the price fluctuations	Central eight tenths of the price fluctuations	Whole number of the price fluctuations
					1st tenth, 27.0					
					2d tenth, 4.9					
					3d tenth, 2.6					
					4th tenth, 2.2					
69.4	11.8	6.1	4.2	3.6	5th tenth, 1.8	3.6	7.8	13.9	25.7	95.1
					6th tenth, 1.8					
					7th tenth, 2.0					
					8th tenth, 3.5					
					9th tenth, 6.9					
					10th tenth, 42.4					

"The central division of the table shows that the average range covered by the fluctuations diminishes rapidly as we pass from the cases of greatest fall toward the cases of little change, and then increases still more rapidly as we go onward to the cases of greatest rise. The right-hand group of columns shows how the range increases if we start with the two middle tenths, take in the two tenths just outside them, then the two tenths outside the latter, and so on until we have included the whole body of fluctuations. The left-hand group of columns, on the other hand, combines in succession the two tenths on the outer boundaries, then the two tenths immediately inside them, and so on until we get back again to the two central tenths. Perhaps the most striking single result brought out by this table is that eight tenths of all the fluctuations are concentrated within a range (25.7 per cent) slightly narrower than that covered by the single tenth that represents the heaviest declines (27 per cent), and much narrower than that covered by the single tenth that represents the greatest advances (42.4 per cent)." *Op. cit.*, p. 17.

The actual distribution of the changes for the 5578 cases is given in the accompanying table, and is compared with a "normal curve of error" in Plate 21.

TABLE B

DISTRIBUTION OF 5578 CASES OF CHANGE IN THE WHOLESALE PRICES OF COMMODITIES FROM ONE YEAR TO THE NEXT, ACCORDING TO THE MAGNITUDE AND DIRECTION OF THE CHANGES (Based upon the chain relative to Table 11 of Bulletin of the Bureau of Labor Statistics, No. 149)

Rising Prices						Falling Prices		
Per Cent of Change from the Average Price of the Preceding Year	Number of Cases	Proportion of Cases	Per Cent of Change from the Average Price of the Preceding Year	Number of Cases	Proportion of Cases	Per Cent of Change from the Average Price of the Preceding Year	Number of Cases	Proportion of Cases
102–103.9	1	0.018	46–47.9	11	0.197	Under 2	[1] 405	7.261
100–101.9	1	.018	44–45.9	10	.179	2– 3.9	[1] 375	6.723
98– 99.9	—	—	42–43.9	6	.108	4– 5.9	329	5.898
96– 97.9	—	—	40–41.9	14	.251	6– 7.9	[1] 238	4.267
94– 95.9	—	—	38–39.9	17	.305	8– 9.9	200	3.585
92– 93.9	—	—	36–37.9	11	.197	10–11.9	173	3.101
90– 91.9	—	—	34–35.9	18	.323	12–13.9	[1] 120	2.151
88– 89.9	—	—	32–33.9	17	.305	14–15.9	107	1.918
86– 87.9	1	.018	30–31.9	22	.394	16–17.9	76	1.362
84– 85.9	1	.018	28–29.9	30	.538	18–19.9	71	1.273
82– 83.9	1	.018	26–27.9	29	.520	20–21.9	45	.807
80– 81.9	1	.018	24–25.9	47	.843	22–23.9	39	.699
78– 79.9	—	—	22–23.9	45	.807	24–25.9	32	.574
76– 77.9	—	—	20–21.9	65	1.165	26–27.9	17	.305
74– 75.9	1	.018	18–19.9	73	1.308	28–29.9	27	.484
72– 73.9	4	.072	16–17.9	[1] 102	1.828	30–31.9	16	.287
70– 71.9	1	.018	14–15.9	106	1.900	32–33.9	7	.125
68– 69.9	3	.054	12–13.9	115	2.062	34–35.9	10	.179

[1] Location of the decils.

Rising Prices						Falling Prices		
Per Cent of Change from the Average Price of the Preceding Year	Number of Cases	Proportion of Cases	Per Cent of Change from the Average Price of the Preceding Year	Number of Cases	Proportion of Cases	Per Cent of Change from the Average Price of the Preceding Year	Number of Cases	Proportion of Cases
66– 67.9	4	.072	10–11.9	167	2.994	36–37.9	7	.125
64– 65.9	—	—	8– 9.9	[1] 237	4.249	38–39.9	5	.090
62– 63.9	—	—	6– 7.9	261	4.679	40–41.9	5	.090
60– 61.9	4	.072	4– 5.9	[1] 356	6.382	42–43.9	4	.072
58– 59.9	6	.108	2– 3.9	355	6.364	44–45.9	2	.036
56– 57.9	1	.018	Under 2	[1] 410	7.350	46–47.9	1	.018
54– 55.9	3	.054	—	—	—	48–49.9	1	.018
52– 53.9	4	.072	No change	[1] 697	12.494	50–51.9	1	.018
50– 51.9	1	.018	—	—	—	52–53.9	—	—
48– 49.9	5	.090	—	—	—	54–55.9	1	.018

Summary

	Number of Cases	Proportion of Cases
Rising prices	2,567	46.021
No change	697	12.494
Falling prices	2,314	41.485
Total	5,578	100.000[2]

In commenting on the distribution and the comparison with the normal error curve, Mitchell says:

"There are three significant points to notice here: (1) The two forms of distribution, the actual and the 'normal,' are of the same type. (2) The concentration about the central tendency is greater in the actual than in the 'normal' distribution; but on the other

[1] Location of the decils. [2] *Op. cit.*, p. 19.

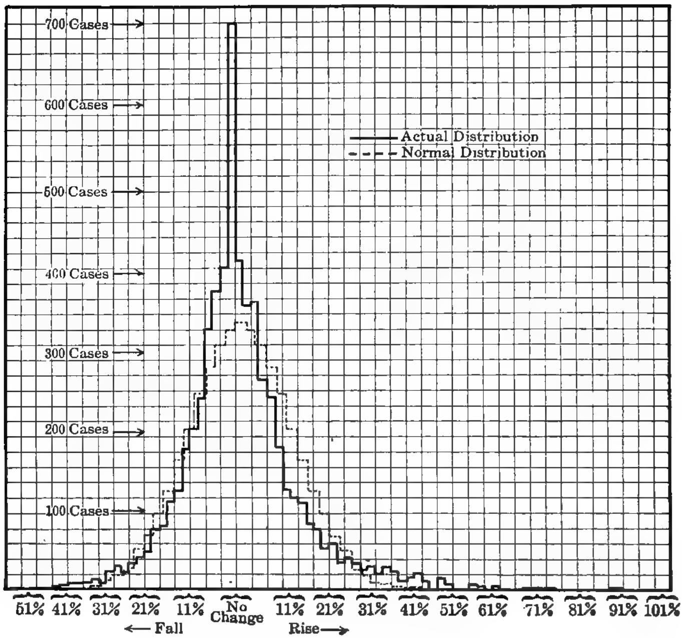


PLATE 21

Distribution of 5578 Price Variations. (Percentages of Rise or Fall from Prices of Preceding Year)

hand, the extreme variations diverge further from this central tendency in the actual distribution than in the other. (3) Unlike the 'normal' distribution, the actual distribution is not perfectly symmetrical. Two closely related aspects of this difference may be pointed out: First, the outlying cases of the 'normal' distribution extend precisely the same distance from the central tendency in both directions, whereas in the actual distribution the outlying cases run much farther to the right (in the direction of a rise in prices) than to the left (in the direction of a fall). Second, the central tendency itself is free from ambiguity in one case but not in the other. In the 'normal' distribution this tendency may be expressed differently by the median, the arithmetic mean, or the mode (the point of greatest density); for these three averages coincide. In the actual distribution, on the contrary, these averages differ slightly; the median and mode stand at ± 0, while the arithmetic mean is + 1.36 per cent. These departures of the actual distribution from perfect symmetry possess a certain significance; but, after all, they are minor qualifications of the important proposition; namely, year-to-year price fluctuations are grouped about their central tendency in a strikingly regular fashion."[1]

The meaning of the agreement between the variations of prices from their normal tendency and the curve of error is important in the interpretation of index numbers. Most numbers, as was said above, are averages of relatives. An average is a summary expression which in and of itself need not reveal the deviations of actual data from an average. These may be large or small and arranged about an average in any form. However, for a normal distribution the variations assume definite form and the median, the mode, and the arithmetic mean agree. Change in price level is then best indicated by an average which subscribes to these conditions. If price indexes are computed in terms of changes from year to year — that is, if chain-relatives are computed — this agreement exists. If they are computed

[1] *Op. cit.*, pp. 19–21.

upon a remote base, the variations do not follow this normal order, and an average of the changes is an imperfect picture of the combined result. Mitchell has stated his conclusion in respect to this point as follows:

"The consequence is that the measurement of price fluctuations becomes difficult in proportion to the length of time during which the variations to be measured have continued. In other words, the farther apart are the dates for which prices are compared, the wider is the margin of error to which index numbers are subject, the greater the discrepancies likely to appear between index numbers made by different investigators, the wider the divergencies between the averages and the individual variations from which they are computed, and the larger the body of data required to give confidence in the representative value of the results." [1]

Two questions of vital interest are raised by the above discussion: First, should reliance be placed in an average of relatives index number? and, Second, if a relative is used what average should be employed? These questions are discussed immediately below.

(2) The Base in Computing a Price Index Number

It has been felt necessary to reduce actual prices to a relative basis in order to combine them. The units in which they are quoted, and the varying importances which are assigned to them, have been in the past quite enough to prevent any reliance being placed in a simple aggregate of the prices of a group of commodities. Absolute differences have been dispelled by the simple expedient of reducing prices of commodities at one period into percentages of the prices which the same commodities bore at another or base period, and by taking the arithmetic or some other average of the aggregate per cents. The result became the index for

[1] *Op. cit.*, p. 23.

the time used. It will be noted, however, that this process nominally amounts to giving all commodities the same weight — that is, unity, since each is called 100 per cent. To correct this, weights have been assigned by arbitrarily giving some commodities more importance than others or by choosing a larger number of those which it is intended heavily to weight. Recently, however, there has developed a tendency to use simply a sum of actual prices, to convert these to a common basis, such as value per pound, and to weight them according to some outward index.[1] By so doing, it is maintained, two difficulties are overcome: First, the problem of choosing a base year, since actual prices do not necessarily have to be reduced to a relative basis, and, second, of deciding upon an appropriate average of relatives.

In the discussion in the preceding section reasons were given for preferring a recent as contrasted with a remote base. The case, however, is not wholly in favor of the use of a recent year or of a chain-relative, although it is no doubt true that most people desire to make comparisons with recent dates, and that year-to-year variations are more accurately measured by an average than are the variations growing out of the use of a remote period. Chain-relatives are difficult to use. Differences from year to year are admirably shown, but not the changes for a period of years.[2] On the other

[1] How generally this is now being done will be seen in the following chapter.

[2] "Of course, chain relatives for successive years can be multiplied together to form a continuous series, but it is not easy to give the later members of the series a concrete meaning. To know, for example, that in 1891 prices fell, on the average, 0.2 per cent below their level in 1890; that in 1892 they fell 4.4 per cent below their new level in 1891, and so on through ups and downs on an ever-changing base for every year to 1915, enables one to make a series beginning, say, with 100 in 1890 and running on with 99.8 in 1891, 95.4 in 1892, etc., to some result for 1915. But such a series does not enable one to say in terms of what a comparison is made between prices in 1915 and in 1890." *Op. cit.*, p. 38.

hand, the ease with which obsolescent commodities may be dropped and new ones added,[1] when actual prices are used; and the further fact that prices with which comparisons are made are recent and do not have to be thought of as "normal" nor "abnormal," but only as actual, are factors tending to increase the popularity and use of the year-to-year type. To change to a new base in the case of an average of relatives requires that the index be re-computed from the beginning, or that the so-called short method[2] be employed. The latter gives doubtful results[2] while the former is prohibitive

[1] "A further advantage of chain index numbers is that they make the dropping of obsolescent and the adding of new commodities especially easy. It is difficult to keep the list of commodities included in a fixed-base system really representative of the markets over a long period of time. Barring perhaps thirty or so staple raw materials that hold their importance for centuries at a time, most commodities have their day of favor and then yield to new products. Consequently the compilers can hardly let two decades pass without revising their lists, in certain details, or seeing them lose in significance. But since a chain index does not profess to give accurate comparisons except between successive years the compiler feels himself free to improve his list whenever he can. It is very much easier to include many commodities on this plan. And if the index number be weighted, the chain index has a similar advantage in facilitating the frequent revision of the weights." *Op. cit.*, p. 37.

[2] "This method consists in dividing the figures for other dates by the figures for the date desired as base and multiplying the quotients by 100. Of course this process results in a relative price of 100 for the new base period, and the other figures look as if they showed average relative prices as percentages of prices at this period. But there is no mathematical justification for assuming that results reached in this way must agree with results reached by recomputing relative prices for each commodity on the new base. For such recomputation usually alters considerably the relative influence exercised upon the arithmetic means by the price fluctuations of certain commodities. Those articles which are cheaper in the new than in the old base period get higher relative prices and therefore increased influence. Vice versa, articles that are dearer in the new base period get lower relative prices and therefore diminished influence. Of course the short method of shifting the base, which retains the old relative prices, does not permit any such alteration in the influence exercised by the fluctuations of different commodities. Hence the two methods of shifting the base seldom yield precisely the same results. To present a series of arithmetic means shifted by the short method as showing what the index numbers would have been if they had been computed upon the new base is therefore misleading." *Op. cit.*, p. 39.

because of the amount of labor involved. When an index is a sum of dollars and cents, it can be put in the form of a relative on any base by a simple numerical calculation.

(3) The Average to Use in Computing a Price Index Number

The discussion of the best average to use in the case of an index number of relative prices has been long and voluminous.[1] It has generally been associated with some phase of the interpretation of price phenomena and has assumed both a mathematical and economic turn. Champions of the arithmetic mean, of the median, and of the geometric mean have appeared. It is not our purpose to enter this discussion further than to call attention to the properties, already discussed, of the more common averages, and briefly to summarize the case for the geometric mean in connection with index numbers.

Some sort of an average is generally used, the most common undoubtedly being the arithmetic mean. Indeed, some have insisted that it is the "natural"[2] average, all others being inappropriate for index number purposes. Others, of which Jevons, the English economist, and Walsh[3] are probably the foremost champions, have insisted upon the geometric mean — that is, the *n*th root of the product of the factors. The merits of any average must of necessity turn upon the nature of the inquiry which is being made. This truth has been so admirably stated by Mitchell in respect to index numbers, that in spite of the emphasis that has already

[1] For instance, see Laughlin, J. L., *The Principles of Money*, Ch. VI, and Bibliography; Mitchell, *op. cit.*, pp. 88 ff.; Fisher, Irving, *The Purchasing Power of Money*, Ch. X, and Appendix to Ch. X; Walsh, C. M., *The Measurement of General Exchange Value*, *passim*.

[2] Padan, *Journal of Political Economy*, 1900, pp. 73 ff., quoted by Laughlin, *op. cit.*, p. 148.

[3] See note 1 above.

been given to it in an earlier chapter, we cannot do better than to quote him.

"Wise choice of the average to use in making an index number, then, involves careful consideration of the materials to be dealt with and of the purpose in view. (1) If that purpose be to measure the *average ratio of change* in prices, the geometric mean is the best, indeed, in strictness, it is the only proper average to employ. For, alone among our averages, the geometric mean always allows equal influence to equal ratios of change in price, quite irrespective of the previous levels of the prices in question, the amounts of money represented by the changes themselves, or any other factor. As has been said already, in a geometric mean the doubling of one price is precisely offset by the halving of another price — though if the two prices were originally the same the rise amounts in money to twice the fall. And further changes of 10 per cent from the two new prices will again be precisely equal in their influence upon a geometric mean, although 10 per cent of the price that has doubled represents a sum of money four times as great as 10 per cent of the price that has been halved. (2) But these same examples show that geometric means are not proper averages for measuring alterations in the amount of money that goods cost. And as a rule our interest does center in the money cost of goods rather than in the average ratio of changes in price. For example, when we are investigating the increased cost of living, the doubling of one item in the family budget may well be twice as important as its halving; and when we are studying the 'relation of prices to the currency, a large upward variation should count for more than a small downward variation, for it requires more currency.' For such purposes the arithmetic mean is the logical average to use. (3) Frequently, however, the very fact that an article has advanced greatly in price cuts down its market, so that the increase in money cost represented by the arithmetic mean exists on paper rather than in fact. When such cases of extreme advance are numerous among the relative prices to be averaged, the median may give more significant results than the arithmetic mean. (4) When the number of commodities included in the index number is small, however, medians are likely to prove highly erratic, representing less the general trend of prices than the peculiarities of the data from which they are made. (5) If

the index number is designed for the public at large, the familiarity of arithmetic means is an argument in their favor; but it counts for nothing in the case of figures intended for specialists. (6) Often the usefulness of a new index number may be enhanced without detriment to its special purpose by throwing it into a form directly comparable with that of index numbers already in existence. Then, of course, not only the form of average but also the base period employed in making the existing series has special claims for imitation. (7) Finally, the desirability of making index numbers that can be shifted from one base to another deserves far more consideration than is commonly accorded it. On this count the score is in favor of the geometric mean. If geometric means were invariably used, all index numbers could readily be compared with one another, whatever the bases on which they were originally computed. And that would be a great gain to all students of prices." [1]

The fact that the geometric mean as an index number can be shifted from one base to another easily and accurately undoubtedly is of advantage.[2] But it is unfamiliar and laborious to compute and is not in general use. It is doubtful if its merits are sufficient to overbalance these last two counts. Certainly not for the general student and business man.

If exceptional changes — these variations far removed from the norm — are to be given weight, and if money costs and their effects are to be taken cognizance of, then the arithmetic mean must be employed so long as averages of relatives are used. But when relatives are calculated upon a remote base, exceptional deviations tend to be exaggerated, the distribution being asymmetrical and not well balanced on either side of the norm. In this respect, so far as both commodity and stock prices are concerned, "geometric means are more significant averages of price fluctuations

[1] Mitchell, *op. cit.*, pp. 88–90.

[2] See the illustration given in Mitchell, *op. cit.*, p. 82.

. . . than arithmetic means, because they are the averages of more symmetrical distributions." [1]

The median also has its champions. Its ease of calculation and the fact that it serves, with the quartiles or decils, to give a notion of distribution of variations about a central tendency cause it to be supported by many. Its characteristics have already been indicated in an earlier chapter, and, following Mitchell, can briefly be summarized in connection with the use in question.

"(1) They are not perfectly reversible; that is, they cannot always be shifted from one base to another by simple division. (2) The median may not answer precisely to its definition when several of the items to be averaged have identical values. . . (3) Medians of different groups cannot be combined, averaged or otherwise manipulated with ease as can arithmetic means. . . . (4) When the number of items to be averaged is small, medians are erratic in their behavior. . . ." [2]

While the virtue of an average is always a function of the use which is to be made of it,[3] this fact is too often ignored in the case of index numbers. Consumers of statistics too readily absorb the completed numbers without bothering themselves over the manner in which they are computed. From this point of view as well as from others, it would be a decided step in advance if index numbers could be computed without resorting to averages at all. This is now done in several cases. However, it is probably a vain hope to hold

[1] Mitchell has made an elaborate comparison of the median, the arithmetic mean, and the geometric mean for stock and commodity index numbers in "A Critique of Index Numbers of Prices of Stocks" in *The Journal of Political Economy*, July, 1916, pp. 625–693. Comparisons of medians and arithmetic means are made in *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173, pp. 87–90.

[2] Mitchell, *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173, pp. 84–85.

[3] This point of view has been developed in Chapter VIII, above, *Averages as Types*.

out that a simplicity of statistical method can ever compensate for blind indifference on the part of user of statistics. More particularly is this true respecting the use of index numbers.

(4) Weighting and its Problems in Connection with a Price Index Number

Distinction is generally drawn between "simple" and "weighted" index numbers. By a weighted number is meant one in which commodities are influential according to their relative importance. When commodities are allowed to influence the result in the same proportion, the result is said to be a "simple" index number. Weighting is effected in various ways. For retail price indexes a common method is to weight according to consumption as revealed in budgetary studies or by aggregate national expenditure. For wholesale price indexes, commodities may be assigned different importances by a conscious choice of the commodities used. In some cases an external index of importance is employed for wholesale numbers, as, for instance, the amount of imports and exports, the amount of production, the value of articles or services "exchanged at base prices in the year whose level of prices it is desired to find." [1] Mitchell has used as weights, in the case of stock index numbers, stock outstanding, earnings, and number of shares sold.[2]

Lack of attention to weights does not mean that weights are equal, but generally that they are haphazard. They are not necessarily bad because of this, nor good, as Mitchell points out, if they are consciously made. "The real problem

[1] Fisher, Irving, *The Purchasing Power of Money*, pp. 217–218.

[2] Mitchell, Wesley C., "A Critique of Index Numbers of the Prices of Stocks," in *The Journal of Political Economy*, July, 1916, pp. 632 ff.

for the maker of index numbers is whether he shall leave weighting to chance or seek to rationalize it." [1]

Moreover, so-called simple index numbers may in fact be markedly weighted; as, for instance, the Aldrich index number, where 25 different varieties of pocket knives were included, thus "giving this trifling article an influence upon the result more than eight times greater than given to wheat, corn, and coal put together." [2] In fact to give each commodity equal weight would require careful and studied attention to the choosing of positive weights.

But what test or tests of importance are available? Are they applicable at all times and places, and for all purposes? If there is in reality no defensible "general purpose" index number, there is likewise no single system of weights of universal application. To weight a retail price index number, where the purpose of its computation is patently to measure the effect of price change on consumers, by the amount of production or by the value of the articles exchanged is ill fitting. Likewise, to weight a wholesale index number, knowing the discrepancies between wholesale and retail prices, by statistics of family budgets is illogical. Reason and fitness must characterize the use of weights — and these must be tested in terms of uses — or they must be dispensed with entirely.

On the relation of weights to purposes of index numbers, Mitchell says:

> "If rational weighting is worth striving after, then, by what criterion shall the relative importance of the different commodities be judged? That depends upon the object of the investigation. If, for example, the aim be to measure changes in the cost of living, and the data be retail quotations of consumers' commodities, then

[1] *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173, p. 72. [2] *Ibid.*, p. 71.

the proportionate expenditures upon the different articles as represented by collections of family budgets make appropriate weights. If the aim be to study changes in the money incomes of farmers, then the data should be 'farm prices,' the list of commodities should be limited to farm products, and the weights should be proportionate to the monetary receipts from the several products. If the aim be to construct a business barometer, the data should be prices from the most representative wholesale markets, the list should be confined to commodities whose prices are most sensitive to changes in business prospects and least liable to change from other causes, and the weights may logically be adjusted to the relative importance of the commodities as objects of investment. If the aim be merely to find the differences of price fluctuation characteristic of dissimilar groups of commodities, or to study the influence of gold production or the issue of irredeemable paper money upon the way in which prices change, it may be appropriate to give identical weights to all the commodities. If, on the other hand, the aim be to make a general-purpose index number of wholesale prices, the question is less easy to answer." [1]

But why use weights at all, when weighted results are so strikingly the same as unweighted? Two main reasons are usually assigned for ignoring them. The first has already been mentioned in the following form: What is the test or tests of importance and where are data to measure it? The second, and one which is thought to be important, is that unweighted series are almost identical with the weighted. Bowley says, in much quoted passages:

"The discussion of the proper weight to be used . . . has occupied a space in statistical literature out of all proportion to its significance, for it may be said at once that no great importance need be attached to the special choice of weights; one of the most convenient facts of statistical theory is that, given certain conditions, the same result is obtained whatever logical system of weights is applied." [2]

[1] *Op. cit.*, pp. 75–76.
[2] Bowley, A. L., *Elements of Statistics*, 2d Ed., 1902, p. 113.

"So we arrive at a very important precept; *in calculating averages give all care to making the items free from bias, and do not strain after exactness in weighting.*"[1]

Weighting properly considered is nothing but a striving after a proper distribution of samples. Sampling may as effectively be done by an adjustment of weights as by the more direct, but sometimes more difficult, method of increasing the commodities taken. In reality the two are alternatives, with this difference that errors in prices will probably tend more nearly to be compensating than those in weights. If a rational system of weight does not change the result of an unweighted average, it may safely be concluded that the latter accurately represents the true condition. If it does, then it may be concluded that the unweighted data are not representative, and that by using weights the effect has been to extend the base so as to include more commodities.

While the problem of selecting weights lends itself to theoretical discussion, it is primarily of practical concern. To the person who desires to use index numbers the question cannot be dismissed with the assertion that if weights are chosen according to chance, weighted and unweighted indexes closely agree. As they are computed, weights are not always so chosen, numbers differ materially, and the merits of unweighted and weighted numbers can be determined only by comparison.[2] In the light of the differences shown in this manner the merits of the two types of series must be determined. The student and business man cannot readily make these comparisons for themselves but they can be familiar

[1] Bowley, A. L., *Elements of Statistics*, 2d Ed., 1902, p. 118.

[2] Weighted and unweighted series, and those weighted in various ways both for commodities and stocks, are elaborately compared by Mitchell, Wesley C., in "Critique of Index Numbers of Prices of Stock," in *The Journal of Political Economy*, July, 1916, *passim;* and *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173, pp. 74–75.

with those that have been made, and can use the indexes in a candid and intelligent manner. That "amiable weakness to take upon faith plausible figures that fill a pressing want" would not then be so common.

Should weights be fixed or fluctuating? By changing them a more accurate measure of importance is undoubtedly acquired, but changes in an index must then be interpreted not only in terms of prices but also in terms of weights. Conceivably, some sort of an average of relative importance over a period could be used, but if so the variations would be lost sight of. When chain-indexes are used, weights can be varied without confusion, since price changes from year to year only are measured. Such figures do not accurately measure changes over a period. The question cannot be answered in a word, and we shall not attempt to settle it. There is much to be said for the stability resulting from the use of fixed weights, and in actual practice necessity frequently requires that one be satisfied with such.

(5) Average of Relatives Index Numbers versus Actual Prices Aggregated

In the section devoted to *The Base* the question of the desirability of actual instead of relative prices was raised, and some of the reasons indicated which have prompted a return to the former kind. This problem may now be considered a little more fully. Two major questions are involved: First, how to reduce commodities quoted in widely different units and in different quantities to a common denominator in order that they can be combined — for price level would not be reflected in the change of a single commodity; and second, what system of weights to use. The first question until recently seemed insuperable. As the Bureau of Labor Statistics puts it:

. . . "it would be a statistical absurdity to make index numbers for the different years from the yearly averages of the actual money prices of a ton of coal, a yard of calico, a hundredweight of live hogs, 144 boxes of matches, a pound of raw rubber, a gallon of turpentine, 50 square feet of window glass, a dozen cans of salmon, a barrel of petroleum, a yard of trouserings, a mule, a pair of boots, a bushel of beans, a thousand feet of pine lumber, a crosscut saw, a barrel of cement, a two-bushel bag, a thousand bricks, a ton of steel rails, a dozen teacups and a dozen saucers, a spool of thread, a pine door, a pound of cotton, a dozen cans of tomatoes, a pair of door knobs, a hundredweight of barbed wire, a hammer, a quintal of codfish, a 'set' of bedroom furniture, a ton of brimstone, a dozen eggs, an apothecary's ounce of quinine, a barrel of salt, a dozen kitchen chairs, a pound of beef, a pair of cotton blankets, a nest of three oak-grained tubs, 100 pounds of onions, a carving set, a bushel of potatoes, a dozen pairs of socks, a three-quarter-inch auger, a barrel of herrings, a troy ounce of silver, a box of raisins, a ton of hay, a dozen undershirts, a quart of milk, a thousand shingles, a yard of broadcloth, a ton of cotton-seed meal, a gross of wood screws, and a pound of plug tobacco." [1]

Even to reduce the various units with the prices quoted per length, dozen, cubical contents, area, weight, etc., to prices per pound, or some other single unit, will not suffice. Left in this manner an index

"greatly exaggerates the effects of price changes in the rare, costly, and relatively unimportant articles, like opium and silver, and correspondingly minimizes the importance of price changes in common, cheap, and important articles, like coal, petroleum, and pig iron. It avoids the inaccuracies of the average of relatives by committing much graver inaccuracies." [2]

To remedy this defect, however, the device is now adopted by the United States Bureau of Labor Statistics in the case of wholesale prices of weighting the price per pound of commodi-

[1] *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 181, *Wholesale Prices*, p. 245.

[2] *Ibid.*, p. 246.

ties by the amount of physical product placed on the market in 1909. In this way a relative of weighted aggregate money prices is secured — the last completed year being the base adopted — instead of an average of relative prices. The theory upon which the number is computed is that "what is wanted in wholesale-price indexes as well as in retail-price indexes is a measure for changes in the cost of a given bill of goods." [1] This purpose seems to be the one in which most people are interested and the sum of actual prices appears best fitted to establish it. Mitchell, after summarizing the advantages of aggregates of actual prices, has the following to say: "Now the weighted aggregate of prices is the best measure of change in the money cost of goods; it is better in several ways than the simple arithmetic mean of relative prices, and in addition it has all the merits of the latter form of average." [2]

"Aggregates of money prices weighted according to the importance of the several articles are as easy to understand as arithmetic means of relative prices. They are less laborious to compute than any other form of weighted series, for no relative prices are used; the original quotations are multiplied directly by the physical quantities used as weights, and the products added together. They are not tied to a single base period; but from them relative prices can quickly be made upon the chain system or any fixed base that is desired, and these relative prices themselves can be shifted about at will as readily as geometric means. Hence they are capable of giving direct comparisons between prices on any two dates in which an investigator happens to be interested. Hence, also, they can be compared with any index numbers covering the same years, on whatever base the latter are computed. Their meaning is perfectly definite — which is not always true of medians. They can not be made to give apparently inconsistent results like

[1] "Wholesale Prices," *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 181, p. 246.

[2] *Ibid.*, Whole Number 173, p. 92.

arithmetic means. When published as sums of money, they can be added, subtracted, multiplied, divided, or averaged in any way that is convenient. When weighted on a sound system, they can not be unduly distorted by a very great advance in the price of a few articles, and yet, unlike medians, they allow every change in the price of every article to influence the result. In fact, they combine most of the merits and few of the defects characteristic of the various methods of averaging relative prices." [1]

IV. Conclusion

The discussion has been carried far enough to establish the fact that index number making and using are far from simple things. The complexity of the problem seemed to make it necessary to develop the various points in this chapter in order to bring before the reader the theoretical and practical considerations surrounding the topic. In most respects little more has been done than to call attention to the more important phases of the subject and to leave the student to verify them by reference to such painstaking and comprehensive studies as those of Fisher, Mitchell, and others. Some of the more important practical applications of the subject are outlined in the following chapter. The aim here is not a critique, but rather an exposition of the principles upon which a critique must be based. If an interest in index number making and using has been aroused, the main purpose of what has been written here shall have been accomplished. After all, the main reliance must be placed in the scientific spirit and integrity of both maker and user. If these are lacking, the use of statistics is without a logical defense.

[1] "Wholesale Prices," *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173, p. 91.

References

Bowley, A. L. — *Elements of Statistics*, pp. 111–118; 217–229.

Coats, R. H. — *Special Report on Wholesale Prices in Canada, 1890–1909*, inclusive (Ottawa, Can., 1910). This report contains a useful summary of the principal index numbers now compiled in the United States and in foreign countries.

Fisher, Irving — *The Purchasing Power of Money*, Ch. 10.

Hooker, R. H. — "The Course of Prices At Home and Abroad, 1890–1910," in *The Journal of the Royal Statistical Society*, Vol. LXXV, pp. 1–36 (December, 1911).

Mitchell, W. C. — *Business Cycles*, pp. 112–139, on "The Representative Character of Index Numbers."

—— — "Index Numbers of Wholesale Prices in the United States and Foreign Countries," United States Department of Labor, *Bulletin Bureau of Labor Statistics*, Whole Number 173, July, 1915. Part I, "The Making and Using of Index Numbers," pp. 5–114; Part II, "Index Numbers of Wholesale Prices in the United States and Foreign Countries," pp. 115–324. (This publication in the field of index numbers is epoch-making. It includes a complete study of the technique of the construction and use, as well as a descriptive account of current and past index numbers in this and in foreign countries.)

Meeker, Royal — "Some Features of the Statistical Work of the Bureau of Labor Statistics," in *The Quarterly Publications of the American Statistical Association*, March, 1915, pp. 431–441.

CHAPTER X

AMERICAN PRICE INDEX NUMBERS DESCRIBED AND COMPARED

I. Introduction

In the preceding chapter the chief considerations in the computation and use of index numbers have been outlined. In this chapter evidence is furnished of the importance of these in the descriptions and comparisons of the leading American index numbers. The treatment is for the most part descriptive, the aim being to emphasize those features which should be known when index numbers are used. The facts here collected, while generally available, are not, it is feared, fully appreciated either by students or by business men. It is with this thought in mind, and with the purpose of giving the theoretical points practical application that a chapter is devoted to the descriptive side of the question.

II. Description of American Index Numbers

American index numbers divide themselves into two groups. First, those currently prepared by the United States Government, and second, those prepared by private establishments. The government issues both a wholesale and a retail number; those published privately are restricted to wholesale prices. The government, moreover, publishes index numbers of wages and hours of labor in certain industries, but a description of these is not included here, inasmuch as the methods are in the main the same as those followed in the price series.

1. *Price Indexes Prepared by the United States Government*

(1) Index of Wholesale Prices Prepared by the United States Government

The systematic publication of a wholesale price index number by the United States Government was begun in 1902. The period first covered was 1890 to 1901, inclusive. This number was in continuation of the index compiled by the Department of Labor for the period 1890 to 1899, but included somewhat different commodities and carried the computation back to 1890. Since then an index has been published annually. Up to and including 1913, the index was an average of relatives based upon the average price 1890–1899. In 1914 a change was made to an aggregate of actual prices, reduced to a price per pound basis and weighted according to the amount of goods placed on the market in 1909. A description of the precise method by which the change was made is deferred until the conditions existing in 1913 have been outlined.

There were 252 commodities included in the index for 1913. The number varied as follows over the period 1890–1913:

TABLE A

TABLE SHOWING THE NUMBER OF COMMODITIES, BUREAU OF LABOR WHOLESALE PRICE INDEX NUMBER, 1890 TO 1913, INCLUSIVE

NUMBER OF COMMODITIES		YEARS	
251	257	1890, 1891	1909–1911
252	258	1913	1906–1908
253	259	1892	1895, 1904, 1905
255	260	1893, 1912	1896, 1899–1903
256	261	1894	1897, 1898

The choice has been such as to give weight to the commodities deemed most important. No definite numerical system of weights was used until the change was made to actual prices in 1914. Before this date the commodities were distributed in groups as follows:

TABLE B

TABLE SHOWING THE NUMBER AND GROUPING OF COMMODITIES FOR THE UNITED STATES WHOLESALE PRICE INDEX, 1890–1913

COMMODITY GROUP	NUMBER	YEARS
Farm products	16	1890–1907
Farm products	20	1908–1913
Foods	53	1890–1892, 1904–1907
Foods	54	1893–1903, 1913
Foods	55	1912
Foods	57	1908–1911
Cloths and clothing	63	1913
Cloths and clothing	65	1909–1912
Cloths and clothing	66	1908
Cloths and clothing	70	1890, 1891
Cloths and clothing	72	1892
Cloths and clothing	73	1893, 1894
Cloths and clothing	75	1895, 1896, 1906, 1907
Cloths and clothing	76	1897–1905
Fuel and lighting	13	1890–1913
Metals and implements	37	1890–1893
Metals and implements	38	1894, 1895, 1899–1913
Metals and implements	39	1896–1898
Lumber and building material	26	1890–1894
Lumber and building material	27	1895–1907
Lumber and building material	28	1908–1913
Drugs and chemicals	[illegible]	1890–1913
House furnishing goods	14	1890–1913
Miscellaneous	13	1890–1913

TABLE C

Table Showing the Number of Commodities or Series of Quotations Classified by Markets for which Prices were Secured, 1913. United States Bureau of Labor Wholesale Price Index Number

Markets	Total	Farm Products	Food, etc.	Cloths and Clothing	Fuel and Lighting	Metals and Implements	Lumber and Building Materials	Drugs and Chemicals	House Furnishing Goods	Miscellaneous
Total	252	20	54	63	13	38	28	9	14	13
New York . . .	129	3	45	2	9	21	23	9	6	11
Chicago	22	14	6	—	—	1	1	—	—	—
Factory, mine, etc. .	11	—	—	—	3	1	3	—	3	1
Pittsburgh	7	—	—	—	—	7	—	—	—	—
Philadelphia	4	—	—	—	—	4	—	—	—	—
Boston	1	—	1	—	—	—	—	—	—	—
Trenton, N. J.	3	—	—	—	—	—	—	—	3	—
Cincinnati	2	—	—	—	1	1	—	—	—	—
Eastern Market . . .	2	—	—	2	—	—	—	—	—	—
East St. Louis	1	1	—	—	—	—	—	—	—	—
Elgin, Ill.	1	—	1	—	—	—	—	—	—	—
LaSalle, Ill.	1	—	—	—	—	1	—	—	—	—
Louisville, Ky.	1	1	—	—	—	—	—	—	—	—
Peoria, Ill.	1	—	—	—	—	—	—	—	—	1
Minneapolis	1	1	—	—	—	—	—	—	—	—
Washington, D. C. . .	1	—	1	—	—	—	—	—	—	—
Wilmington, N. C. . .	1	—	—	—	—	—	1	—	—	—
General Market . . .	63	—	—	59	—	2	—	—	2	—

In 1913, of the 252 commodities, 45 were "raw" and 207 "manufactured." Over the whole period, 1890 to 1913,

234 identical series were used, and in the last year 44 were weekly prices and 208 monthly. Standard trade journals furnished the price quotations of 129 articles; official boards of trade, 9; chamber of commerce, 1; produce exchanges, 7; leading manufactures, 105; and a government bureau, 1 article. New York market furnished the price quotations for 129 articles; Chicago, for 22; "general market," for 63. The remainder were distributed at various points over the country. The distribution of commodities for which prices were secured in 1913, classified by markets, is shown on preceding page.

Numerous changes since the series was begun have been made in the articles included, due to changes in commercial importance, lack of suitable quotations, discontinuance of manufacture, etc. In each case, however, the articles substituted have been as nearly alike those discontinued as was possible. Of a typical change the Bureau says:

"For example, nutmegs were dropped in 1908 because they were insignificant in the economy of the people. The price quotations were dependable, but a rise or fall in the price of nutmegs had no importance. . . . In 1904 Danish cloth was substituted for alpaca, and in 1907 Sicilian cloth was substituted for Danish cloth, in order to represent the kind of women's dress goods most in demand at these different periods of time. Eleven new commodities were added to the list in 1908, 2 of which have since been discontinued, while 90 additional price series have been included in the present bulletin to give a fairer and more complete idea of price fluctuations." [1]

The manner of incorporating new commodities into the new index is described by the Bureau as follows:

". . . For example, the prices of Burbank potatoes were quoted down to 1907. In that year the description of potatoes was expanded to include all kinds of white potatoes, good to fancy in

[1] *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 181, pp. 240–241.

grade, thus securing more dependable quotations throughout the year, because some variety of white potato is certain to be in market at all times, while the supply of Burbank potatoes may be very scant or fail entirely. There was no material difference in the price of the two descriptions of potatoes, so it was not necessary to resort to the process of substituting the quotations of potatoes, white, good to fancy, for Burbank potatoes. When a new article differing in quality enough to show a considerable difference in price has been introduced in the place of an article which has become obsolete or which is no longer representative, the prices of the new article have been substituted for the prices of the article dropped in the manner described below. For example, in 1904 Danish cloth at \$0.1125 per yard was substituted for alpaca at \$0.0764 per yard. The average price of alpaca for 1890–1899 was \$0.0680, therefore its relative price in 1904 was 112.4, *i.e.* $\frac{\$0.0764}{\$0.0680} = 112.4$. This relative price of alpaca in 1904 was taken to represent the relative price of Danish cloth in 1904. In 1905 the money price of Danish cloth was \$0.1150. This money price was reduced to a relative price for 1905 on the 1904 price as a base, giving $\frac{\$0.1150}{\$0.1125} = 102.2$. This 1905 relative price of Danish cloth calculated on its 1904 price as a base was then multiplied by the 1904 relative price of alpaca on the 1890–1899 average price as the base in order to shift the 1905 relative price to the 1890–1899 base. This operation gives 114.9 ($102.2 \times 112.4 = 114.9$) as the relative price of Danish cloth in 1905. . . ." [1]

This method of substitution was followed when prices of the original and the substitute goods could be gotten for the same year. When such prices were not available, a different method was pursued, change being made necessary by the addition in 1908 of 11 new commodities. The method has been severely criticized,[2] and it is pretty certain that a

[1] *Ibid.*, pp. 242–243.

[2] See Mitchell, Wesley C., *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173, July, 1915, pp. 42–44.

realization of the justice of criticism[1] was a potent reason for the Bureau's change to an aggregate of actual prices. Concerning the change the Bureau says:

"The method adopted by the bureau may best be made clear by describing how the index number of a particular group was computed. Let us consider the farm products group. In this group horses, mules, live poultry, and Burley tobacco were included for the first time in 1908. Prices of these new articles were obtained for both 1907 and 1908. A relative price for each of the 20 old and new articles included in the group was calculated for 1908 on the 1907 base. These relative prices were added together and divided by 20, the number of commodities in the group, to get the simple arithmetic average of the relative prices of farm products in 1908 on the 1907 base. This group index number was then multiplied by the 1907 index number computed on the money prices of the 16 old articles to obtain the 1908 index number of farm products on the 1890–1899 base. . ."[2]

The uncertainty of this method, the difficulty of changing an average of relatives computed on a remote base to a recent one without entirely recomputing the series,[3] and the realization that a relative price "built up from actual money prices shows much more accurately what we want to show,

[1] Meeker, Royal, "Some Features of the Statistical Work of the United States Bureau of Labor Statistics," *Publications of the American Statistical Association*, March, 1915, pp. 431–442.

[2] *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 181, p. 244.

[3] The limitations of the "short method," notice of which was made in the last chapter, are acknowledged in the following words by the present Commissioner of Labor Statistics: "A more 'scientific' method employed is to divide both relative prices through by the 1912 relative. . . . The Bureau has resorted to this method in previous bulletins, to construct tables purporting to show the percentage changes in prices from year to year. This method of procedure is mathematically unsound and the result is vitiated by an amount of error that can be ascertained only by digging up the original price data and reconstructing the relative prices anew on the 1912 base." Meeker, Royal, "Some Features of the Statistical Work of the United States Bureau of Labor Statistics," *Publications of the American Statistical Association*, March, 1915, p. 439.

namely, change in the cost of living, — changes in the cost of the same quantity of a commodity or of an unvarying market basket," [1] — resulted in the Bureau's change to aggregate actual money prices.

Beginning in 1914, for wholesale prices, the Bureau changed to this basis.[2] Briefly the changes were as follows: Forty-one distinct articles were dropped, 31 new ones were added, while the number of quotations was increased by including prices from all of the larger cities where acceptable ones were available. "These changes were necessary in order to make the list represent more accurately the bulk of commodities exchanged and the great markets where exchanges are effected at wholesale in the United States at the present time." [3]

The base period was shifted from the average of prices for the ten-year period, 1890–1899, to the last completed year; in this case, 1914. Two reasons for so doing were assigned by the Bureau.

". . . this change was made for the purpose, first, of utilizing the latest and most trustworthy price quotations as the base from which price fluctuations are to be measured, and second, to permit of the addition of new articles to those formerly included in the index number. For practically all articles which it was desired to add to the index no prices were obtainable for the period 1890–1899." [4]

The method of making the shift is described by the Bureau as follows:

"The price of each article in 1914, the base year, has first been multiplied by the quantity of the article marketed in the last census year, 1909. The products thus obtained have then been summed, giving the approximate value in exchange in 1914 of all articles in

[1] *Ibid.*, p. 436.

[2] Details are shown in *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 181, October, 1915.

[3] *Ibid.*, p. 5.

[4] *Ibid.*, p. 5.

the group or in the total list of commodities. Similar aggregates have likewise been computed for each year from 1890 to 1913 and for each month of 1913 and 1914. With the aggregate for 1914 as the base, or 100, the index number for each year prior to 1914 and for each month of 1913 and 1914 has been obtained by comparing the aggregate value for such year or month with that for 1914. . . ."[1]

By using the farm products group, the precise method may be illustrated. The aggregate value of this group in 1914 (the sum of the average price of each article in 1914 multiplied by the quantities of each marketed in 1909) was \$4,334,063. This was taken as 100. The aggregate for the same commodities in 1913 was \$4,191,601. This divided by the 1914 aggregate equals 96.7 and gives the index for 1913. The aggregate in 1912 was \$4,224,483. For identical articles in 1913 the aggregate was \$4,187,367, and stood in relation to the 1912 aggregate as 100 to 100.9. The index for 1912 was obtained by multiplying the index for 1912 on the 1913 base (100.9) by the index for 1913 on the 1914 base (96.7), *i.e.* 100.9×96.7. This gave a product of 97.6, the index for 1912 on the 1914 base.[2]

The Bureau now publishes four wholesale series, two major or primary ones and two that are derivative. The first is the unweighted average of relatives based upon the average price 1890–1899 and continues the series which dates back to 1890. The second is the weighted aggregate of actual prices. The other two are derived from these. Just how far these are comparable is an open question which the Bureau itself does not answer. It does, however, call attention to the inherent difference and warns against hasty comparisons.[3]

An important feature of the Bureau's work is the publica-

[1] Details are shown in *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 181, October, 1915, p. 6.

[2] The details with figures are contained in *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 181, pp. 257–263.

[3] *Ibid.*, pp. 10–11.

tion, along with the index numbers, of the actual prices of the commodities used. These constitute the raw material for special and independent studies.

In brief, the index number of wholesale prices published by the United States Bureau of Labor Statistics is now a weighted aggregate of actual prices, reduced to a relative basis. It is computed on the basis of 340 commodities and seems to be designed for the purposes of measuring changes in the cost of a quantity of commodities, not particularly to the consumer, nor the producer, not to the investor, nor the speculator, but to any of these. As such it is a general-purpose number, made up from prices of raw and manufactured commodities, consumers' and producers' goods, including forest and animal products, drawn from the larger cities and industrial centers.

The weights assigned are the quantities of the goods marketed in 1909 — the last date for which adequate statistics are available. Just what the change to actual prices will mean in the nature of the series, it is probably too early definitely to say. It may, however, positively be asserted that the Bureau is thoroughly converted to the wisdom of the change, since it has been extended to all of the series which the Bureau issues. Certainly the candid manner in which the problem of change has been met and the illuminating discussion by the Bureau of the reasons for and the effects of the change cannot but be reassuring. One feels that the change has been made in good faith, that the occasion demanded it and that the new plan has been worked out in a scientific manner.

(2) Indexes of Retail Prices Prepared by the United States Government

If the collection of price data as a basis for the computation of a wholesale price index presents real problems, as it undoubtedly does, these are many times more serious in the case of price data for a retail price index. While retail prices may change more slowly than wholesale, may be less affected by trade disturbances, and may move further in either direction after they are disturbed and be slower to regain their former position, it is these conditions and others, which make it so difficult to procure satisfactory price data over a period of time so as to measure the changes actually taking place. Prices of some commodities change from day to day; others less susceptible to conditions of demand and supply show appreciable change within somewhat longer periods. Prices for the same commodity vary materially as between localities. Some commodities, standard in character, but peculiar to local markets and not possessing distinctive trade names, sell at widely different prices at the same time. If the problem is to measure price level for retail prices, the commodities to be chosen, the frequency with which quotations are to be taken, and the regions from which prices are to be collected are serious questions. These and others discussed in Chapter IX must be settled before the collection of actual prices is begun. Because so many questions of technique in the collection and so many principles of method in the handling of data are involved in computing retail price index numbers, it is thought advisable fully to describe the methods employed by the United States Bureau of Labor Statistics.

The Bureau's retail price index is avowedly a consumer's number. Only materials which enter into the budget of a

typical American workingman's family are included, and prices are taken from industrial centers. The weights applied vary according to the proportions in which commodities enter into such a budget.

From 1890 to 1907, 30 commodities were used. From 1907 to 1913, this number was reduced to 15, and in 1914 and 1915, respectively, the number was 17 and 21. The additions were made possible because of the Bureau's change in 1914 from an average of relatives to an aggregate of actual prices. Price data were received (1915) from 725 dealers, 150 bakeries, 215 retail coal dealers, 65 gas companies, and 205 dry goods stores, located in 45 industrial centers in 34 states. The base price from 1890 to 1913 was the average for the ten-year period, 1890–1899; since 1913 it has been the last completed year.

The detail of the method employed by the Bureau from 1890 to date may be summarized as follows[1]:

a. The Period 1890–1903, Inclusive

Identical firms quoted prices during the complete period. A yearly relative for each of the thirty commodities was computed on the base, average of the prices, 1890–1899. Relatives for each commodity for the various firms reporting in a city were added and the sum divided by the number of reporting firms to get the city relative. City relatives for each commodity within each of the geographical divisions chosen by the Bureau for the presentation of data were added and the sum divided by the number of geographical divisions to get a divisional relative. Likewise, the city

[1] A detailed account, upon which the following is based, of the change made in 1914, in computing the United States Bureau of Labor Statistics Retail Price Index, is given in *Bulletin* of the Bureau, Whole Number 156, pp. 357–380.

relatives for each commodity were added and the sum divided by the number of cities to get a relative for the country at large. An average of all the relatives taken in this form furnished an index of the price level for the country.

b. The Period 1904–1907, Inclusive

Changes in the firms reporting in 1904 made it necessary to devise some method of incorporating their prices into the index. The method chosen was as follows: All new firms furnished prices both for 1903 and 1904. For each commodity, the 1904 price was put in the form of a relative on the 1903 base; these relatives were added and the simple average taken, as above described, for indexes for cities, for geographical divisions, and for the country as a whole. To convert each commodity to a relative on the 1890–1899 base, the 1904 relative on the 1903 base was multiplied by the 1903 relative on the 1890–1899 base.

"For example, in the North Atlantic division it was found that the average relative price of wheat flour in 1904 as compared with its average price in 1903 was 117.91. The average relative price of wheat flour in 1903 as compared with its average price for the period 1890–1899 was 101.6. Multiplying 117.91 by 101.6 gives 119.8, which was taken as the relative price of wheat flour for this geographical division in 1904 on the 1890–1899 base." [1]

c. The Period 1908–1913, Inclusive

Beginning with 1908, 15 commodities were dropped from the index "because the quality of some of the articles changed so radically from year to year and even from month to month." [2] "The method of computing relative prices

[1] *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 156, p. 359.
[2] *Ibid.*, p. 359.

employed from 1903 to 1911, inclusive, involved computing the relatives on the preceding year as the base, and afterwards shifting to the 1890–1899 base by multiplying by the relative for the preceding year computed on the 1890–1899 base." [1]

However, beginning with 1912, because of the failure of many firms to report regularly, and because of the omission of some of the commodities, it was decided to compare identical firms month by month. This was thought to be necessary because price changes can be compared accurately only by including identical firms and identical articles. The method required that a relative for each commodity for each firm be computed for both December, 1911, and January, 1912. These relatives were then added and an average taken of the firms in the cities for city relatives, and the city relatives combined and averaged to get a divisional index and an index for the country as a whole. The relative prices for each commodity were then shifted to the 1890–1899 base by multiplying them by the December relatives computed on the 1890–1899 base. The prices reported for the identical firms for January and February were compared by obtaining February relatives on January (as January had been on a December) base and were then shifted to the 1890–1899 base by multiplying through by the January relative computed on the 1890–1899 base. This process was repeated for each month. The yearly relatives for each commodity were obtained by averaging the monthly relatives. The process was followed until January, 1914, at which time the change was made to an aggregate of actual prices.

[1] *Ibid.*, p. 366.

d. The Period 1914 to Date

In accounting for the change to actual prices the Bureau says:

". . . it is apparent that the relative prices of individual commodities, as well as the combined relative prices of all commodities or index numbers, as heretofore constructed, are averages of percentages. The firm relatives were averaged to get the city relative, the city relatives were averaged to get each geographical division relative and also the United States relative. The individual commodity relatives for the country and its divisions were averaged to produce the combined relative or index number for all commodities for the whole country and its divisions; and finally, the monthly relatives were averaged to get the yearly relatives for firms, cities, geographical divisions and the United States.

"When averages of averages of relative prices are thus piled up, it becomes difficult to comprehend the meaning of the final average, even if no theoretical or mathematical errors are involved in the processes.

"A simple arithmetic average of percentages is useful for certain purposes, but for the purposes of retail-price studies which should show changes in expenditures by consumers, a percentage based on average or aggregate *actual* prices of a commodity reflects more accurately the changes in the cost of that commodity." [1]

The difference in the two methods of computing index numbers the Bureau shows in the following manner by taking actual prices of a commodity whose variations are violent and irregular, reported by identical firms in a single city.

An extreme case is taken, as the Bureau says:

"to show that the difference in principle of the two methods of computing relative prices is not of theoretical interest only, but presents quite startling differences in results, which cannot be ignored or set aside with the assertion that 'in the long run' dif-

[1] *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 156, March, 1915, p. 364.

ferences tend to disappear and 'in the end' the results will be approximately the same. Experimentation goes to show that differences in results do not tend to disappear." [1]

TABLE D

TABLE SHOWING DIFFERING RESULTS OBTAINED BY TWO METHODS OF COMPUTING RELATIVE PRICES OF A SINGLE COMMODITY [2]

(Actual prices are for potatoes in Baltimore, Bulletin No. 132, p. 29)

Firm	Actual Price		Relative Price	
	May	June	May	June
A	$.24	$.28	100	116.7
B	.32	.30	100	93.8
C	.24	.28	100	116.7
D	.24	.40	100	166.7
E	.35	.25	100	71.4
Aggregate	1.39	1.51	500	565.3
City relative price	100	108.6	100	113.1

The difference in this case is 4.5 points or more than 4 per cent. In the new method equal *actual* changes in price have the same effect on the result; in the old method equal *percentage* changes have the same effect.[3]

The method of shifting the base when averages of relatives are used, as was the case from 1903 to 1911, inclusive, on a yearly base, and from 1912 to 1913, inclusive, on a monthly base (both described above) is now held by the Bureau to be wrong and to involve —

[1] *Ibid.*, pp. 365–366.
[2] *Ibid.*, p. 365.
[3] *Ibid.*, p. 366.

"an amount of error which is greatest when prices differ most in the base period and change most capriciously from time to time." [1]

The amount of error involved for such a commodity is illustrated by the following table taken from one of the Bureau's reports:

TABLE E

TABLE SHOWING DIFFERING RESULTS OBTAINED BY SHIFTING BASE PERIOD OF RELATIVE PRICES COMPUTED BY OLD AND NEW METHODS

Potatoes. (An example of an article whose prices fluctuate widely and capriciously) [2]

Firm	May		June			July		
	Price	Relative on May base	Price	Relative on May base	Relative on June base	Price	Relative on June base	Relative on May base
804	$0.20	100	$0.40	200.00	100	$0.30	75.00	150.00
808	.17	100	.30	176.47	100	.32	106.67	188.24
815	.50	100	.40	80.00	100	.35	87.50	70.00
817	.20	100	.20	100.00	100	.30	150.00	150.00
821	.20	100	.40	200.00	100	.35	87.50	175.00
City aggregates	1.27	500	1.70	756.47	500	1.62	506.67	733.24
City relatives — averages of firm relatives		100		151.29	100		101.33	146.65
City relatives computed from actual prices, *i.e.* \$1.70 ÷ \$1.27, \$1.62 ÷ \$1.70, \$1.62 ÷ \$1.27		100		133.86			95.29	127.56

[1] *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 156, March, 1915, p. 367.

[2] *Ibid.*, p. 367.

City relative for July on May base computed by averaging relatives and multiplying the averages, *i.e.* 151.29 × 101.33 = ...	153.30
City relative for July on May base computed by multiplying relatives computed from aggregate actual prices, *i.e.* 133.86 × 95.29 = ...	127.56

If the above table is interpreted in terms of the Bureau's old and new methods, the following differences in results are apparent:

(*a*) The relative price for June on the May base, computed from an aggregate of actual prices, is 133.86, *i.e.* \$1.70 (the sum of the actual prices for June) divided by \$1.27 (the sum of the actual prices for May). The similar result for July on the June base is 95.29.

(*b*) The relative price for June on the May base, computed from an average of relatives, is 151.29, *i.e.* $\frac{1}{5}$ of 756.47 (the sum of the June relatives on May). The similar result for July on the May base is 146.65.

(*c*) Shifting the base by the method followed by the Bureau in 1912 and 1913, *i.e.* from month to month, and averaging relatives and multiplying the averages, give the following results:

(*a'*) The June relative on the May base is 151.29.

(*b'*) The July relative on the June base is 101.33.

(*c'*) The July relative on the May base is 151.29 × 101.33, or 153.30, which is 6.65 points greater than 146.65, the result of computing July relative directly on May.

(*d*) Shifting the base by the new method of multiplying relatives computed from actual prices, gives 133.86 × 95.29, or 127.56, as contrasted with 153.30, the result from the old method of shifting. Shifting by the new method can be done "with mathematical accuracy *so long as the actual price quotations come from identical firms throughout the period*

considered."[1] This is undoubtedly a decided advantage of the new over the old method, as indicated in the last chapter.

Base shifting by subtracting the index numbers of commodities at two periods, as for instance, 1912 and 1913, when they are computed by the old method, and calling the difference the percentage of gain, is of course meaningless. Even the more refined method formerly resorted to by the Bureau, of dividing through by the relative for 1912, for instance, is now acknowledged by the Bureau to be wrong and to involve an amount of error which can "be ascertained only by going back to the original actual prices and reconstructing the relative prices anew on the 1912 base."[2] This the Bureau does for two commodities — the difference in the case of potatoes between the correct and the incorrect method being ten points. The Bureau adds:

"This is not an imaginary example, set forth for the purpose of showing a theoretical possibility that contains no element of probability. The example in which the prices of potatoes are used is extreme, but such capricious fluctuations are repeated each year for potatoes and to a certain extent for eggs and such commodities as are subject to violent price changes. Potato prices are used as an example to show typical price changes in a commodity that fluctuates capriciously in price, as the prices of round steak[3] are used to illustrate typical price changes in commodities that fluctuate rather narrowly."[4]

The relative prices computed from actual prices can be shifted to any base without error, the reason being that relative prices are simply ratios of actual prices. "Dividing through by the relative price of any year or period merely has the

[1] *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 156, p. 369.

[2] *Ibid.*, p. 370.

[3] For this commodity the difference is 0.52 point, *Ibid.*, p. 370.

[4] *Ibid.*, p. 371.

effect of substituting the aggregate actual price for the base period as divisor in the formula for computing the relative price."[1] In a final summary of the weakness of the old method, the Bureau says:

> "By the old method of computation any errors which may have existed in price data in the base period 1890–1899 would affect the series of relatives throughout the entire period covered. Errors were introduced by means of the method of averaging relatives calculated from different prices as bases, and these errors were cumulated by the process of shifting the base of the relative prices every month. These inaccuracies taken with the inflexibility of relative prices and indexes calculated by averaging relatives made the changes in methods of calculation which have been carried out imperatively necessary."[2]

The changes of 1914 consisted in adopting the last completed year as a base, and using actual prices from month to month returned by identical firms. The yearly aggregate for 1913 — the base used — was computed by comparing the actual prices reported by identical firms month by month with January, 1913, aggregating these and dividing by 12. How this was done may be illustrated as follows: Eighty-nine identical firms reported prices of granulated sugar for both January and February, 1913. Dividing the aggregate February price by the aggregate January price gave the February relative on a January base. In February and March, 86 identical firms reported prices of this commodity. Dividing the March aggregate price by the aggregated February price gave the March relative on a February base, and multiplying this by the February relative on the January base gave the March relative on the January base. A repetition of this process gave the relatives for each month for each commodity on the January base. The aggregate

[1] *Ibid.*, p. 372. [2] *Ibid.*

of these relatives was then divided by 12 to get the relative for the year. No error was involved in so doing, since all were computed on the same base, viz., 89 firms in January. The base was then shifted from January, 1913, to the average for the year by dividing through by the yearly average calculated on January. In the case of this commodity the yearly relative on January was 94.5, and the monthly relatives on the 1913 base (calculated as above) were January, 105.8; February, 100.0; March, 98.7; etc.

In a similar manner the index number for each commodity, for each geographical division, and for the country as a whole, on the 1913 base, was extended back month by month for the years 1911 to 1913, inclusive, for every second month [1] for the years 1907 to 1910, inclusive, and year by year for the years 1907 to 1913, inclusive.

Such in brief are the old and new methods pursued by the Bureau in computing a retail price index number. But the Bureau, besides showing price indexes, as averages of relatives, for commodities separately, combined these into two series. The first was a simple unweighted number computed by taking the arithmetic average of the sum of relatives of individual commodities. The second was a weighted index in which the relatives for each commodity were weighted according to a scale of consumption based upon the findings of the United States Commissioner of Labor in a study made in 1901 into 2567 workingmen's family budgets.[2] This likewise was an average of relatives, the divisor being not the number of commodities, but the sum of the weights. The method employed in 1913 to get this weighted average is shown in the following table:

[1] From 1907 to 1910 inclusive the Bureau had received prices for every second month only.

[2] *Eighteenth Annual Report of the United States Commissioner of Labor*, Washington, D.C., 1901.

TABLE F

TABLE SHOWING THE WEIGHTS APPLIED TO RELATIVE PRICES TO GET A WEIGHTED INDEX NUMBER

ARTICLES	RELATIVE IMPORTANCE	RELATIVE PRICE	RESULT
Fresh beef	1,531	180.9	276,957.9
Fresh hog products	429	213.8	91,720.2
Salt hog products	425	203.6	86,530.0
Poultry	290	171.8	49,822.0
Eggs	514	174.8	89,847.2
Milk	652	140.2	91,410.4
Butter	880	153.2	134,816.0
Lard	286	166.6	47,647.6
Sugar	482	95.3	45,934.6
Flour and meal	513	138.4	70,999.2
Potatoes	395	151.2	59,724.0
Total	6,397	163.4	1,045,409.1[1]

The divisor in this case is 6397[2] and the dividend 1,045,409.1. The quotient — the index for the year — is 163.4.

When the change was made to an aggregate of actual prices, it would have been meaningless to have combined all the quotations into a single sum. The number of firms reporting and the number of quotations included were not constant factors, and to combine them would have been to make the index depend upon the number of firms and quotations as well as upon the price changes themselves. To avoid this a new method was worked out by the Bureau and followed

[1] *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 156, p. 363.

[2] The combined expenditure is taken to equal 10,000. The commodities used by the Bureau constitute $\frac{6397}{10000}$ of the total.

for the 1914 and 1915 combined retail price indexes. To describe this method, granulated sugar is taken as a typical commodity.

"The aggregate actual price of granulated sugar in January, 1913, for North Atlantic division ($4.9502) was multiplied successively by the relative prices of granulated sugar on the January base for each month of the year 1913. A series of monthly price aggregates was thus built up on the assumption that the 89 stores reporting in January had continued to report throughout the year. The arithmetic average of these aggregates for the 12 months of 1913 was taken as the average aggregate actual price for the year 1913. This average aggregate price ($4.6779) for 1913 was divided by 89, the number of firms reporting in January and the number assumed as reporting throughout the year, to obtain the average actual price of granulated sugar (5.26 cents) for the year 1913. This computed average actual price of granulated sugar in 1913 was next multiplied by the amount of sugar consumed in the North Atlantic division in 1901, according to the Eighteenth Annual Report of the Commissioner of Labor. This formula, $0.0526 × 283 lbs. = $14.89 gives the cost of the amount of sugar consumed by the average workman's family in 1901, purchased at the average price obtaining in 1913. In like manner the cost in 1913 of all other commodities at retail was computed by calculating first the average price of each commodity for 1913 and then multiplying this average price by the quantity consumed in 1901." [1]

Such a combined index is worked out for the years prior to 1913 by aggregating the costs of each of the commodities consumed in 1901, which costs are determined by multiplying the cost of the quantities consumed in 1901 on the basis of 1913 prices by the index number for the earlier years worked out on the basis of 1913, according to the method described above. That is, in the case of granulated sugar, the cost of 283 pounds (the amount consumed according to the study made of workingmen's budgets) in terms of 1913 prices was

[1] *Op. cit.*, p. 377.

found to be \$14.89. This amount is multiplied by 122.6 (the relative price of granulated sugar in January, in this case, on the 1913 base), which gives \$18.26 as the price of 283 pounds of this commodity at the average price in January, 1912. Treating all other commodities in the same manner, and aggregating the costs, they amount to \$328.52—the total cost of a food budget for the North Atlantic division in January, 1912. The cost of the same budget in 1913 prices was \$333.90. Therefore the relative cost of the budget in January, 1912, calculated on the 1913 base was $\frac{\$328.52}{\$333.90}=98.4$. Relative costs for an unvarying budget were computed for each month and for the year 1912 as well as for prior years, and constitute the new retail index for such periods.

This discussion it is feared has been somewhat long and involved. To have fully described the Bureau's methods in all their detail would have taken even more space and probably would have been more involved. For a more complete discussion recourse must be had to other sources.[1] Because of the statistical devices which the old and the new methods illustrate, and more particularly, because of the lack of care with which index numbers however computed are used for any and almost all purposes, it is felt that the discussion has been worth while. The willingness to proceed by averages without at the same time having knowledge of where one is being led could not better be illustrated than in the practices of the Bureau before the recent change. A realization of the weaknesses in the old method finally became so overwhelming that the Bureau was willing to acknowledge its error, to reconstruct its number on a new basis, and to defend

[1] *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 156, March, 1915.

its action in detail. This shows candor and integrity and should be given wide publicity. A study of the change and of the methods involved in making it cannot but help to cause greater reliance to be placed in the Bureau's number, and a better understanding to be had of just what *method* in statistical analysis means in such a case.

2. *Price Indexes Prepared by Private Establishments*

The discussion of price indexes prepared by private establishments will be briefer than that for the government series for the reasons: first, that less is known about them, and second, that the principles of index number making have been fully illustrated in the treatment of the government series. While there are many private series compiled only three — Bradstreet's, Dun's, and the Annalist's — will be discussed. Section III,[1] taken from Professor Mitchell's masterly analysis of index numbers, currently compares seven series — public and private.[2]

(1) Bradstreet's Index Number

Bradstreet's is a wholesale number, based upon 110 to 96 articles, is published monthly in the form of the sum of actual prices of the commodities reduced to a per-pound basis. The articles included are divided into thirteen groups as follows: Breadstuffs, live stock, provisions and groceries, fresh and dried fruits, hides and leather, raw and manufactured textiles, metals, coal and coke, mineral and vegetable oils, naval stores, building materials, chemicals and

[1] pp. 361–376.

[2] For a complete discussion of these and other American series as well as foreign series, see Mitchell, Wesley C., "Index Numbers of Wholesale Prices in the United States and Foreign Countries," *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173, pt. II.

drugs, and miscellaneous. The sum of the different indexes for the 13 groups is the index for the whole number of articles. Yearly indexes are derived by averaging the 12 monthly totals. No base is used and it is not clear from the descriptions contained in Bradstreet's whether the prices used are averages of extremes or something else. Moreover, the source of the quotations is not disclosed. If missing data are interpolated for, neither this fact nor the method employed is published. Weights are not used, except as they appear in the process of reducing all quantities to a price-per-pound basis. This, of course, results in employing a —

> "curious combination of rational and irrational weights. The rational element consists in the inclusion of several quotations for important articles like pig iron, coal, lumber, and hog products, and only one quotation for articles like lemons, tea, and flax. The irrational element results from the reduction of all the original quotations to prices per pound. On April 1, 1897, these prices per pound ranged from $0.0008 for soft coal and coke to $0.52 for quicksilver and $0.83 for rubber. Recognition of the excessive influence upon the results accorded to these high-priced articles presently led the computers to drop them from the index number; but they seem to have retained articles like alcohol and Australian wool which in 1897 cost $0.33 and $0.49 per pound — 400 and 600 times as much as soft coal and coke." [1]

The index is illustrated in the following table, which gives the numbers for the first of January, April, July, and October for each of the years 1907–1914 inclusive:

[1] *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173, p. 101.

TABLE G

TABLE SHOWING BRADSTREET'S INDEX NUMBER FOR SELECTED MONTHS FOR 1907–1914, INCLUSIVE

Year	Bradstreet's Index Number: First of the Month			
	January	April	July	October
1907	$8.9172	$8.9640	$9.0409	$8.8506
1908	8.2949	8.0650	7.8224	8.0139
1909	8.2631	8.3157	8.4573	8.7478
1910	9.2310	9.1996	8.9246	8.9267
1911	8.8361	8.5223	8.5935	8.8065
1912	8.9493	9.0978	9.1119	9.4515
1913	9.4935	9.2976	8.9521	9.1526
1914	8.8857	8.7562	8.6566	9.2416

(2) Dun's Index Number

Dun's index number is based upon the wholesale prices of about 200 commodities from the principal markets of the United States. It is in the form of the amount in dollars and cents required to purchase a year's supply of goods for an individual at the time named. No base, therefore, is necessary. The commodities are divided into seven groups.

"Breadstuffs include quotations of wheat, corn, oats, rye, barley, beans, and peas; meats include live hogs, beef, sheep, and many provisions, lard, tallow, etc.; dairy and garden products embrace eggs, vegetables, fruits, milk, butter, cheese, etc.; other foods include fish, liquors, condiments, sugar, rice, also tobacco, etc.; clothing covers the raw material of each industry, as well as quotations for woolen, cotton, silk, and rubber goods, also hides, leather, and boots and shoes; metals include various quotations for pig iron and partially manufactured and finished products, as well as the minor metals, tin, lead, copper, etc., and coal and petroleum;

miscellaneous includes many grades of hard and soft lumber, lath, brick, lime, glass, turpentine, hemp, linseed oil, paints, fertilizers, and drugs." [1]

The same authority from which the above is quoted gives the following account of the method by which the number is computed:

"Quotations of all the necessaries of life are taken and in each case the price is multiplied by the annual per capita consumption, which precludes any one commodity having more than its proper weight in the aggregate. Thus, wide fluctuations in the price of an article little used do not materially affect the "index," but changes in the great staples have a large influence in advancing or depressing the total. . . . The per capita consumption used to multiply each of many hundreds of commodities does not change. There appears to be much confusion on this point, but it should be seen at a glance that there would be no accurate record of the course of prices if the ratio of consumption changed. It was possible, however, to obtain figures sufficiently accurate to give each commodity its proper importance in the compilation. This was done by taking averages for a period of years when business conditions were normal and every available trade record was utilized, in addition to official statistics of agriculture, foreign commerce, and census returns of manufactures." [2]

The following table shows Dun's numbers for the first of the months, January, April, July, and October, for the period 1907 to 1914, inclusive.

[1] *Dun's Review*, May 9, 1914, quoted by Mitchell, Wesley C., in *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173, p. 150.
[2] *Op. cit.*, p. 149.

TABLE H

TABLE SHOWING DUN'S INDEX NUMBER FOR SELECTED MONTHS FOR 1907–1914, INCLUSIVE

YEAR	DUN'S INDEX NUMBER: FIRST OF THE MONTH			
	January	April	July	October
1907	\$107.264	\$107.895	\$113.660	\$116.140
1908	113.282	108.728	108.174	109.991
1909	111.848	116.864	119.021	118.301
1910	123.434	121.555	119.168	115.449
1911	115.102	110.928	118.130	119.292
1912	123.438	128.049	122.277	123.106
1913	120.832	119.217	116.319	123.902
1914	124.528	119.791	119.708	123.531

(3) The Annalist's Index Number

The *Annalist*, a New York financial journal, publishes weekly an index number based upon the wholesale prices of 25 food products. The commodities are chosen so as to represent the principal items in a family budget. The series dates back to 1913 and is an average of relatives, the base period being the average price of the ten years, 1890–1899. The prices are those of New York and Chicago markets. No weights are used, the method of computing being to take the simple average of the relatives of 25 commodities. Weekly, monthly, and yearly numbers are published.

The following table shows the numbers for the months of January, April, July, and October for the years 1912 to 1914, inclusive:

TABLE I

TABLE SHOWING THE ANNALIST'S INDEX NUMBERS FOR SELECTED MONTHS FOR THE YEARS 1912–1914, INCLUSIVE

YEAR	THE ANNALIST'S INDEX NUMBER FOR			
	JANUARY	APRIL	JULY	OCTOBER
1912	139.681	152.326	143.285	141.861
1913	137.197	141.971	139.839	141.664
1914	142.452	141.120	144.879	150.245

Without attempting further to give a detailed description of American index numbers in current use, the differences between them and the causes for the same may be shown by quoting extensively from a study of Professor Mitchell. Although his comparison includes seven series it admirably suits our purpose. After showing in various ways and by a series of tables the extent of the differences between the numbers considered, Professor Mitchell has the following to say concerning the degree of and causes for the same: [1]

III. COMPARISON OF AMERICAN WHOLESALE PRICE INDEX NUMBERS

"The man who thinks that index numbers do well if they get within 10 per cent of the truth might be satisfied with this showing. But the man who hopes for three significant digits [2] would be disappointed if he had to accept these seven series as similar in meaning and equal in authority. For the detailed differences among them

[1] Mitchell, Wesley C., "Index Numbers of Wholesale Prices in the United States and Foreign Countries," *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173, pp. 98–112.

[2] Or for two significant digits when the index number is less than 100.

are neither few nor trifling. . . . For example, (1) the net change in the price level between 1890 and 1913 is made twice as great by two series as it is made by two others; (2) the maximum difference between any two series for a given year averages over 11 points and varies irregularly between the wide limits of 3 and 19 points; (3) in a year of such decided business character as 1908 two of the series show a rise of 6 to 8 points, while four indicate a fall of 7 to 12 points; (4) indeed the seven series all agree about the direction of price changes in only 12 cases out of 23; (5) regarding the degree of these changes from one year to the next they show discrepancies ranging all the way from 2 to 20 points and averaging nearly 10 points for the whole period; (6) the seven series also differ strikingly in respect to steadiness, the least steady making the average change in prices from one year to the next almost twice as great as the steadiest series makes it; (7) certain of the series reflect changes in business conditions with marked regularity, others are quite unreliable business barometers, etc.

"To show that these series differ in many details, however, means little. The significant problem is whether these differences are due to the inherent difficulty of measuring changes in the price level, to the crudity of the general method of measurement in vogue, or to technical differences in the construction of the particular index numbers in question. . . .

"The seven series may be analyzed with respect to the ultimate sources of information drawn upon, the adequacy of the original quotations of each commodity, the numbers and kinds of commodities included, the weights employed, the use made of relative prices, and the kinds of average struck. At each step the question is whether the observed differences among the index numbers accord with the differences found to be characteristic of the various methods considered. If most of the differences can be accounted for in this way, considerable confidence may be felt in the possibility of measuring approximately the variations in prices by index numbers.

"The sources of information, the frequency of the quotations, and the forms of average used are in part so little known and in part so similar that they give us no help in explaining the discrepancies among the results. On the contrary, a marked influence can be traced with confidence to differences in methods of weighting and in the numbers and kinds of commodities included.

"Dun's index number is said to be weighted by per capita consumption, and the weights for the separate commodities are so arranged that foods count for 50 per cent of the total, textiles for 18 per cent, minerals for 16 per cent, and other commodities for 16 per cent. Gibson's index number in its present form is also said by the publisher to be weighted according to Dun's method.[1] . . .

"Haphazard weighting preponderates also in the two series from the Bureau of Labor Statistics, for the representation accorded to different commodities has not been thoroughly worked out on any logical plan. It is true that in the original figures certain highly important articles are represented by two or more series — for instance, coal, iron, cattle, and leather; but so also are certain articles of slight moment, such as window glass, glassware, saws, sheetings, etc. In the two remaining index numbers, the Annalist series and the original form of Gibson's index number, no formal weights are applied; but the lists of commodities have been carefully studied and the most important articles allotted two or three sets of quotations.

"The constitution of the seven series with respect to the numbers and kinds of commodities included can best be represented in tabular form. The analysis, given in the next table, can not be applied to Dun's index number for lack of information about the commodities and weights used, and it can not be strictly applied to Gibson's present series because we know the commodities but not the weights allotted each. In the case of Bradstreet's index number the percentages of the total are computed on the basis of the prices per pound of 96 commodities published for April 1, 1897. This basis is not wholly satisfactory, because the relative price per pound of different commodities, and therefore their relative influence upon the result, has doubtless changed considerably from year to year. But the error arising from using these figures for a single date is less than the error that would arise if we merely counted the number of Bradstreet's commodities in the several classes.[2] In dealing

[1] For Mitchell's criticism of the weights used by Bradstreet's, see *supra*, p. 357.

[2] "Bradstreet's now publishes quotations of 106 commodities, bases its index number on quotations of 96, and does not tell which 10 are omitted. Its prices per pound, published for only a short while in 1897, include 98 articles, among them rubber and quicksilver, which are known to have been dropped from the index number at a later date. Accordingly the quota-

with the remaining series counting the number of commodities in each class is satisfactory, since there are no weights to be considered aside from the number of forms or products by which each article is represented.

TABLE J

Analysis of the Commodities Included in the Leading American Index Numbers

1. Division into raw, slightly manufactured, and manufactured products.

Index Number	Total Number of Commodities	Number of Commodities Classified as			Percentage of Total		
		Raw	Slightly Manufactured	Manufactured	Raw	Slightly Manufactured	Manufactured
1. Bureau of Labor Statistics, original	242	49	25	168	20	10	70
2. Bureau of Labor Statistics, revised	145	36	21	88	25	14	61
3. Bradstreet's	96	40	22	34	[1] 36	[1] 9	[1] 55
4. Gibson, original	50	26	4	20	52	8	40
5. Annalist	25	8	5	12	32	20	48
6. Gibson, present	22	11	2	9	50	9	41

tions for the remaining 96 articles have been accepted as the basis of this analysis. Their prices per pound sum up to $5.9154, whereas Bradstreet's revised index number for this date is $6.0460 — a difference of about 2 per cent."

[1] Percentage of the total weights on April 1, 1897, not of the number of commodities included.

2. Subdivision of the manufactured and slightly manufactured goods.

Index Number	Total Number of Commodities	Number of Commodities Classified as			Percentage of the Total		
		Consumers' Goods	Producers' Goods	Both Consumers' and Producers' Goods	Consumers' Goods	Producers' Goods	Both Consumers' and Producers' Goods
1. Bureau of Labor Statistics, original	193	108	73	12	45	30	5
2. Bureau of Labor Statistics, revised	109	51	47	11	35	32	8
3. Bradstreet's	56	21	30	5	[1]26	[1]26	[1]12
4. Gibson, original	24	11	12	1	22	24	2
5. Annalist	17	17	—	—	68	—	—
6. Gibson, present	11	11	—	—	50	—	—

3. Subdivision of the raw materials and slightly manufactured goods.

Index Number	Number of Commodities	Number of these Commodities Classified as				Percentage of the Total			
		Farm Crops	Animal Products	Forest Products	Mineral Products	Farm Crops	Animal Products	Forest Products	Mineral Products
1. Bureau of Labor Statistics, original	74	18	15	12	29	7	6	5	12
2. Bureau of Labor Statistics, revised	57	18	10	10	19	12	7	7	13
3. Bradstreet's	62	24	15	6	17	[1]14	[1]25	[1]1	[1]5
4. Gibson, original	30	10	8	3	9	20	16	6	18
5. Annalist	13	6	7	—	—	24	28	—	—
6. Gibson, present	13	8	5	—	—	36	23	—	—

[1] See Note, p. 364.

"What light do these facts about weights and the numbers and kinds of commodities included shed upon the differences among the seven index numbers?

"To begin with, the present Gibson and the Annalist index numbers are confined to one kind of commodities — foods, or rather foods and the staples from which foods are prepared. The other index numbers include besides foods an equal or greater number of textile materials and fabrics, minerals, building materials, fuels, drugs, etc. The constitution of the seven series in this respect is as follows: [1]

Index Number	Whole Number of Commodities	Number of Foods	Per Cent of Foods
1. Bureau of Labor Statistics, original	242	58	24
2. Bureau of Labor Statistics, revised	145	40	28
3. Bradstreet's	96	37	[2] 29
4. Gibson, original	50	21	42
5. Dun's	310?	?	[2] 50
6. Gibson, present	22	22	100
7. Annalist	25	25	100

"Now it has been shown above that food index numbers differ widely and capriciously from miscellaneous-list index numbers, because the prices of agricultural products are largely dependent upon the yield of each season's harvests, while the prices of most other articles are less dependent upon weather conditions than upon the activity or depression of business. Hence, if index numbers are sufficiently accurate to charge their very differences with meaning, the seven series under analysis should fall into three groups. (1) The two index numbers composed exclusively of foods should resemble each other rather closely and should differ rather widely

[1] Foods are here taken in the rather liberal sense implied by the present Gibson and Annalist index numbers. Hence the number of foods credited to the Bureau of Labor Statistics is greater than the number of articles which it so classifies in its own index number.

[2] Weights allotted foods. Bradstreet's weights as of April 1, 1897.

from the three series in which foods count for less than a third of the total. (2) These three series, in turn, should resemble each other closely and differ, not only from the food indexes pure and simple, but also, though in less measure, from the two series in which foods count for approximately half of the total. (3) The latter, Dun's index number and the index number made from Gibson's original list, should be hybrids, standing intermediate between the two pure stocks, Dun's inclining rather toward the food index numbers and Gibson's toward the miscellaneous-list group.

"These expectations are put to the test in the next table and handsomely realized. The best simple criterion of relationships among the index numbers is the average number of points by which their results differ for each of the 24 years for which data are available. On this basis it appears that the two forms of the Bureau of Labor Statistics' series and Bradstreet's index number come very close together — the greatest average difference is only 2 points. On the other hand, the two food index numbers agree much better with each other than they agree with any of the other series — though the average difference between them is 3.9 points — distinctly larger than the differences among the miscellaneous-list series. Presumably, this greater difference arises from the relatively small number of articles included by both the Annalist and Gibson's present list, 25 and 22, respectively. Finally, it also turns out not only that Dun's index number and the series made from Gibson's original list stand between the two extreme groups, but also that of the two the Gibson series bears a distinctly greater resemblance to the miscellaneous-list group and Dun's index number a rather closer resemblance to the food group." [1]

[1] Note omitted.

TABLE K

DEGREES OF KINSHIP AMONG THE SEVEN AMERICAN INDEX NUMBERS AS SHOWN BY THE AVERAGE NUMBER OF POINTS BY WHICH THEY DIFFER IN THE YEARS 1890 TO 1913

1. Average differences between the original form of the Bureau of Labor Statistics index number and —

	POINTS		POINTS		POINTS
Bureau of Labor Statistics, revised	1.0	Gibson, original	2.5	Annalist .	6.6
Bradstreet's	1.9	Dun's . . .	5.5	Gibson, present form	7.2

2. Average differences between the revised form of the Bureau of Labor Statistics index number and —

	POINTS		POINTS		POINTS
Bureau of Labor Statistics, original .	1.0	Gibson, original	2.0	Annalist . .	6.3
Bradstreet's .	2.0	Dun's . .	5.3	Gibson, present form	6.8

3. Average differences between Bradstreet's index number and —

	POINTS		POINTS		POINTS
Bureau of Labor Statistics, original .	1.9	Gibson, original	3.5	Annalist . .	6.7
Bureau of Labor Statistics, revised .	2.0	Dun's	6.6	Gibson, present form	7.0

4. Average differences between the index number made from Gibson's original list and —

	POINTS		POINTS		POINTS
Bureau of Labor Statistics, original .	2.5	Dun's . .	4.1	Annalist .	5.5
Bureau of Labor Statistics, revised . .	2.0			Gibson, present form	5.9
Bradstreet's . . .	3.5				

5. Average differences between Dun's index number and —

	Points		Points		Points
Bureau of Labor Statistics, original . .	5.5	Gibson, original	4.1	Annalist . .	6.1
Bureau of Labor Statistics, revised .	5.3			Gibson, present form . .	4.5
Bradstreet's . . .	6.6				

6. Average differences between the Annalist index number and —

	Points		Points		Points
Bureau of Labor Statistics, original . .	6.6	Dun's . . .	6.1	Gibson, present form .	3.9
Bureau of Labor Statistics, revised .	6.3	Gibson, original	5.5		
Bradstreet's .	6.7				

7. Average differences between the present form of Gibson's index number and —

	Points		Points		Points
Bureau of Labor Statistics, original . .	7.2	Dun's . . .	4.5	Annalist . .	3.9
Bureau of Labor Statistics, revised . .	6.8	Gibson, original	5.9		
Bradstreet's . .	7.0				

"Gibson's present series, then, and the Annalist index number may be set aside as different in kind from the miscellaneous-list series. They do not aim to measure the same thing as the latter, and therefore the wide and frequent discrepancies between the two groups are not disquieting. Quite the contrary, the series differ from the miscellaneous-list series in precisely the ways that the previous sections would lead one to expect. This fact is highly reassuring;

for it means that in different parts of the business field there really are general trends among the apparently random variations of prices, and that existing index numbers have measured these divergent trends with approximate accuracy. Otherwise such close consistency would hardly exist among the results.

"It is equally reassuring to find that most of the small discrepancies among the three miscellaneous-list series are also consistent with what has already been learned about the price fluctuations of different kinds of commodities. Indeed it is curious that two such dissimilar kinds of weighting as are used in Bradstreet's index and in the two series drawn from the Bureau of Labor Statistics should not have produced wide discrepancies. These three series never contradict one another flatly about the direction in which prices are moving. The nearest approach to disagreement occurs in the five years (1893, 1897, 1903, 1904, and 1913) when one or two fail to change while another moves up or down a trifle. In no year are the two bureau series more than 4 points apart, and their average difference is only 1 point. Similarly, Bradstreet's is never more than 7 points out with the original bureau index, and never more than 6 points out with the revised series. Its average differences from them are 1.9 and 2 points, respectively. Bradstreet's is sometimes above and sometimes below the two bureau series, so that its average differences from them computed from algebraic sums of the plus and minus quantities are only five-tenths and nine-tenths of 1 point, respectively. The corresponding average difference between the two bureau series is four-tenths of 1 point.[1]

"The discrepancies that do occur arise chiefly from the fact that while a given change in business conditions affects all three series in the same way it usually causes a wider fluctuation in Bradstreet's index than in the revised bureau series, and a wider fluctuation in the latter than in the bureau's original series. This difference in steadiness is just what should follow from the constitution of these three index numbers with reference to their proportions of raw materials and manufactured products. To the reader who remembers that raw materials fluctuate much more widely in price than goods manufactured from them, the following schedule tells its own story:

[1] It is interesting to compare these differences with those which separate

Index Number	Average Change from Year to Year	Percentage of Raw Materials
	Points	
Bureau of Labor Statistics, original	4.0	20
Bureau of Labor Statistics, revised	4.1	25
Bradstreet's	5.6	36

"The only thing that is difficult to explain, indeed, is the general level on which the three index numbers fluctuate in 1900–1913. We should expect Bradstreet's to stand a little higher than the

the index numbers worked out above for different parts of the system of prices.

Index Number	Difference		
	Average	Maximum	Minimum
Bureau of Labor Statistics, original, and Bureau of Labor Statistics, revised	1.0	4	—
Bureau of Labor Statistics, original, and Bradstreet's	1.9	7	—
Bureau of Labor Statistics, revised, and Bradstreet's	2.0	6	—
49 raw materials and 183 to 193 manufactured articles	5.9	18	1
20 raw materials and 20 of their products	9.1	21	—
5 raw materials and 5 groups of their products	14.0	28	—
Mineral and farm products	10.1	31	—
Mineral and animal products	9.0	32	—
Mineral and forest products	18.6	61	—
Farm and animal products	8.3	20	1
Farm and forest products	19.6	47	1
Animal and forest products	15.8	41	1
Producers' and consumers' goods	6.7	19	1

Note. — For the figures from which these differences are computed see Tables 18, 9, 10, and 11. (Reference is to Professor Mitchell's Tables.)

two bureau indexes because of its larger proportion of raw materials and smaller proportion of minerals. In fact it stands a shade lower, and the slight weight it assigns to the rapidly rising prices of forest products seems hardly sufficient to account for this result, since these products count for only 5 and 7 per cent of the totals in the two bureau series. . . .

CRITICAL VALUATION

"A just evaluation of our seven American index numbers is not easy to make. For a comparison has little meaning unless it deals with all the important points at which the series differ. And since no one series is superior to the others at all points a verdict can not be rendered in a single sentence.

"In the publication of actual prices, the Bureau of Labor Statistics and Bradstreet's stand foremost. The contribution they have thus made to the knowledge of prices possesses great and permanent value over and above the value attaching to their index numbers. For, it is well to repeat, all efforts to improve index numbers, all investigations into the causes and consequences of price fluctuations, and all possibility of making our pecuniary institutions better instruments of public welfare depend for their realization in large measure upon the possession of systematic and long-sustained records of actual prices. And much of this invaluable material would be lost if it were not recorded month by month and year by year.

"Critical users of statistics justly feel greater confidence in figures which they can test than in figures which they must accept upon faith. Hence the compilers of index numbers who do not publish their original quotations inevitably compromise somewhat the reputation of their series. They compromise this reputation still further when they fail to explain in full just what commodities they include, and just what methods of compilation they adopt.[1] In the latter respect the Annalist index number shares first honors with the Bureau of Labor Statistics' series. Any one who chooses to take the trouble can find what commodities are used, and how the final results are worked up from the raw material. Bradstreet's index number suffers a bit in comparison because readers are not told

[1] Note omitted.

which 96 commodities out of the 106 of which prices are published are included in the index number, and because the method of reducing prices by the yard, the dozen, the bushel, the gallon, etc., to prices per pound is not fully explained. Dun's index number is more mysterious still, because neither the list of commodities nor the weights applied to each commodity are disclosed. And Gibson's present series also stands partly in the shadow because, while the list of commodities is known, the publishers state merely that these articles are weighted by Dun's system.

"With reference to weighting, Bradstreet's index number takes low rank, for the plan of reducing all quotations to prices per pound grossly misrepresents the relative importance of many articles. That figures made thus should give results in close agreement with the Bureau of Labor Statistics' series is a remarkable demonstration of the ability of index numbers to extract substantial truth even from unpromising materials. The agreement is all the more remarkable since the bureau's series is also badly weighted, though in a different way and in less degree.[1] The revised bureau series is scarcely better than the original in this respect. It is better in substituting a single set of relatives for the articles of minor importance to which the original accorded several sets (for example, shirtings, sheetings, tools, window glass, etc.), but worse in cutting down the representation accorded to great staples (for example, pork, coal, pig iron, and leather).[2] The Annalist index number follows the sensible, though rudimentary, plan of including two or three varieties of the most important articles, and only one of the less important. The like can be said in favor of Gibson's index number, both in its original and its present form, and in addition Gibson uses the Dun system of weights. The latter system is, in theory, the nearest approach to a satisfactory plan of weighting made by any American index number at present. Whether the practice is as good as the theory is doubtful, to say the least, for any one familiar with the deficiencies of American statistics of consumption must wonder whence the compilers derived their estimates of the quantities of 310 commodities 'annually consumed by each inhabitant.' Moreover, what little is known concerning the actual weights is not unobjectionable. Fifty per cent of the total is too large a weight to allow to foods in a wholesale-price

[1] Note omitted. [2] Notes omitted.

series. Even in the great collection of budgets of workingmen's families made by the Commissioner of Labor in 1901 the average expenditure for food was less than 45 per cent of total family expenditure;[1] and in wholesale markets, of course, many commodities that are never directly consumed by families have great importance.

"Dun's index number is supposed to stand first in number of commodities included, but lack of definite information makes it impossible to judge whether its list is well balanced. The bureau's list also is long and contains samples of many different kinds of goods, manufactured as well as raw, consumed for all sorts of purposes and produced under all sorts of conditions; but the representation accorded to different parts of the whole system of prices is certainly far from equitable. Bradstreet's list, while less than half as long as the bureau's, seems better chosen. It is particularly strong in raw materials and rather weak in manufactured goods. The same remarks apply to Gibson's original list, though it suffers in comparison by being only about half the length of Bradstreet's. Finally, the present Gibson index number and the Annalist series are confined to foodstuffs, and make no pretense of representing prices at large.

"In the form of presenting results, Bradstreet's set an admirable example, which was wisely followed by Dun's. Their sums of actual prices can readily be turned into relatives on any base desired, and hence can be made to yield direct comparisons between any two dates. The other series, as averages of relative prices on the 1890–1899 basis, cannot be properly shifted without a detailed recomputation of the relative prices of each commodity, and force readers to make all their comparisons in terms of what prices were in the decade used as base.

"It is interesting, finally, to test the reliability of the several index numbers as 'business barometers.' Monthly figures would be much better than our yearly averages for this purpose; but since they are not to be had for most of the series during most of the period covered, we must do the best we can with the rougher gauge. In 11 of the 23 cases of changes from one year to the next the seven index numbers disagree as to whether prices rose, fell, or remained constant. In the following schedule these 11 years are represented by columns in which each index number is credited with plus one

[1] Note omitted.

when its change accords with the character of the alteration in business conditions, debited with minus one in cases of disagreement, and marked zero when it recognizes no change in the price level.[1] The net scores made by casting up the plus and minus entries indicate roughly the relative faithfulness with which these series have reflected changes in business conditions in the past. Of the index numbers regularly published, Bradstreet's makes much the best showing. Even the scores against it in 1895 and 1903, and its failure to show the reaction in business conditions in 1913, would be wiped out were the data by quarters and months used in place of the annual averages.

Index Number	1891	1893	1895	1897	1901	1903	1904	1905	1908	1910	1913	Net Score
1. Bradstreet's	[2] + 1	+ 1	− 1	+ 1	+ 1	− 1	+ 1	+ 1	+ 1	+ 1	0	+ 6
2. Bureau of Labor Statistics, revised	+ 1	+ 1	− 1	0	+ 1	0	0	+ 1	+ 1	+ 1	+ 1	+ 6
3. Gibson, original	0	0	0	+ 1	+ 1	+ 1	− 1	+ 1	+ 1	+ 1	0	+ 5
4. Bureau of Labor Statistics, original	+ 1	0	− 1	0	+ 1	− 1	+ 1	+ 1	+ 1	+ 1	0	+ 4
5. Annalist	− 1	− 1	− 1	+ 1	− 1	+ 1	− 1	+ 1	− 1	+ 1	+ 1	− 1
6. Dun's	− 1	− 1	− 1	− 1	− 1	+ 1	− 1	0	+ 1	+ 1	+ 1	− 2
7. Gibson, present	− 1	− 1	− 1	+ 1	− 1	+ 1	+ 1	− 1	− 1	0	+ 1	− 2

"Each of these seven series, then, has its special uses, its merits, and its defects. Choice among them should be made in accordance with the particular purpose for which an index number happens to be wanted. But it seems feasible to construct an American series which would present a stronger combination of good qualities as a general-purpose index number than any now existing. The original quotations might be collected from the records of the Bureau of Labor Statistics and Bradstreet's, a list of commodities more complete than Bradstreet's and better balanced than the bureau's might be drawn up, the use of actual prices might be adopted from Brad-

[1] For a description of American business conditions in this period, see W. C. Mitchell, *Business Cycles*, Chapter III (Summary, p. 88).

[2] Based on Bradstreet's original figures for 1890 and 1891, figures which are not used in the index number as currently published.

street's and Dun's, the several commodities might be weighted by physical quantities after Dun's fashion, but with the use of a criterion more appropriate to wholesale prices, and the whole process of construction might be set forth with the frankness characteristic of the Annalist and the bureau. Such a series might differ little from the figures now available; but, however it might turn out, its results would merit greater confidence than can properly be felt in any of the present index numbers as a measure of changes in the general level of wholesale prices."

IV. Conclusion

The collection of data, the development of plan and purpose, the use of statistical abbreviations in the forms of averages and aggregates, the association of means and ends are all admirably illustrated in index number making and using. With few statistical problems is it necessary to use so many data and to exercise so much care in the uses to which they are put, and yet these facts are not generally acknowledged by those who use index numbers and are likely to be given little weight unless the consequences of loose and indiscriminate use are pointed out. It has been the purpose of this part of the discussion briefly to develop the principles of index number making and to show their importance in respect to the leading American numbers. The application of statistical method is patent at every stage.

CHAPTER XI

DESCRIPTION AND SUMMARIZATION — DISPERSION AND SKEWNESS

I. Introduction

Perhaps it is well at this time to restate the order of our treatment. It proceeds from the simple to the complex; from detail to summary. Statistical data are first to be collected; they are then to be dissected and appraised for the purpose in mind, and afterwards to be combined into aggregates for comparative purposes. Comparison may be of time or place, of extent or condition, but in all statistical work it is the goal.

Averages have been treated as summarizing expressions.[1] They seem to bring to focus in a single expression the dissimilarities and peculiarities of data. How inadequate they sometimes are, however, in this respect is apparent from the differences which frequently exist between them, and from the further fact that in matters of social interest — wherein a norm or "average" is unreal or does not exist — deviations

[1] Pearl, in speaking of the functions of statistics, says that they give us "Knowledge of certain *abstract qualities* of groups or masses. This . . . is obtained by calculation from the counted data." These important qualities are: (*a*) "The center or typical condition" — giving the mean, median or mode; (*b*) "The degree of undividual diversity," giving the average and the standard deviations; and (3) "Degree of symmetry." This knowledge is exact "*so long as we confine our attention solely to the particular group discussed in a particular single case.*" For example: Average heights to the nearest inch of three men would not give a "reasonable" measure if they were widely different. Pearl, Raymond: *Modes of Research in Genetics*, pp. 80–81.

or differences from an average are far more important than an average itself. The "reality" of such summaries is much less certain in the fields of economic and social statistics than it is in natural science, where according to an orderly arrangement, excesses and deficiencies above and below a characteristic thing, in respect to a given phenomenon, arrange themselves about a norm in a predictable manner. Not infrequently even a few samples if properly chosen perfectly reflect this natural order. Not that averages in economics are of no use; quite the contrary. They clearly have a function, but it is too frequently abused by not being properly understood. Their precision in the field of natural science is too frequently blurred and obscured when they are applied in business and economics. They still give impressions and roughly characterize statistical distributions, but rough characterizations and general impressions are inadequate as bases for important social, business, and economic changes. It is detail that must somehow be incorporated, but not so as to confuse the issue in its larger aspects. The problem of the statistician is to make data vivid in outline and at the same time to incorporate within them essential detail. Moreover, these must be apparent and be given proper weight.

The logic of large numbers is not forgotten in this connection. It has already been recognized that one need not have complete statistical data on all phases of an economic problem in order to understand it. Statistical sampling is so general as almost to be characteristic. Sometimes it is followed because of choice but more frequently perhaps because of necessity. But there is a vast difference between arriving at a conclusion from adequate statistical samples and of stating this conclusion solely by means of statistical abbreviations. It is the latter which is now being considered.

While employing averages as statistical abbreviations it is possible to supplement them in such a way that details will not be sacrificed, much less be ignored. By the use of simple measures of dispersion and skewness, definite meaning may often be given to facts which, if expressed by averages alone, would be inadequate and possibly misleading in all cases where discrimination is important. It is to a description of these that this chapter is devoted. It is thought best not only to explain the function of such measures but fully to illustrate, by the means of concrete examples, their application to economic statistics as well as the methods by which they are calculated.

II. Dispersion

1. *The Meaning of Dispersion*

Dispersion is the term used to express the variability or difference of the separate measures in a group (frequency series) or in a time series from the average or characteristic feature. Dispersion calls attention to the degree of homogeneity which characterizes statistical groups. If the limits established are wide, as they are, when nothing more respecting a loan, for instance, is noted than the fact that it is a loan, the rates of interest are widely different. That is, the dispersion or "scatteration" from the average is large. On the other hand, if municipal loans for a single purpose are compared, the range of difference between the interest rates is noticeably narrower. That is, the dispersion is smaller.

Of course, highly dissimilar things can hardly be said to have a characteristic feature, and to be described by a *single* expression. Difference and variation are characteristic of most things. Absolute uniformity, rarely found in natural, is not even approached in many economic phenomena.

Freight cars differ as to capacity; engines, as to tractive power; people, as to earning capacity; etc. It is the differences or variations from the characteristic thing which it is the function of measures of dispersion to reveal.

In matters of pure chance and in natural phenomena, frequencies tend to be distributed about a norm or central tendency in a regular and orderly way. The error or normal law of error curve is described. Median, mode, and arithmetic mean coincide. Distribution is symmetrical irrespective of the types of the series. They may be discrete or continuous. The fact of symmetry, however, does not reveal the amount by which the variables are more or less than the average or typical fact. They may be small as in Plate 21, Chapter IX, or large as in Plate 20, Chapter IX. It is these which measures of dispersion reveal. Averages alone are inadequate; comparisons of them are enlightening.

2. *Measures and Coefficients of Dispersion*

It is of advantage to distinguish between time and frequency series when treating measures of dispersion. In time series the controlling fact is chronology; in frequency series, amount. This fact makes the treatment of the two somewhat different.

(1) The Range.

The limits of a distribution or series may be established by citing the range within which frequencies fall. In frequency groupings the units are cited; in time series the upper and lower limits of distributions are given. Extremes in the latter case, however, need not correspond to the time limits, since arrangement is according to chronology and not amount. This is illustrated in the time series shown in Chapter VIII.

According to the arrangement on p. 269, the minimum and maximum amounts, 46,631,000 and 121,852,000, respectively, do not coincide with the time limits of the period. To express the limits of the amounts is to ignore the limits of the period, and *vice versa*. The arrangement follows the order of amount, and violates that of time. This is necessary for the determination of the median and quartiles, but is not common in tabulation.

On the other hand, in frequency series, when extremes of amounts are listed, minimal frequencies usually correspond. This is always true in symmetrical curves and is approached in those which are moderately asymmetrical. Maximal frequencies, on the other hand, correspond in normal distributions to normal amounts, and approach the same in moderately asymmetrical ones. Merely to express the range, however, may mean very little in either case. Light is not necessarily thrown on the nature of the distribution between the extremes. In historical series they may almost be coincident in point of time. In frequency series, they may mean very little because they are unrepresentative. These facts are further considered by use of examples.

In the series on p. 269, Chapter VIII, the extremes are 46,631,000 lbs. and 121,852,000 lbs. But these alone tell nothing concerning the distribution between the limits. Certainly the minimum is far more characteristic of the series than is the maximum. The extremes would not be altered by a very different order. Again, using the frequency distribution in Table M, Chapter VIII, the extremes are $5.00 to $5.99 and $14.00 to $14.99, but the frequencies for the minimum are fifteen times as large as those for the maximum. Something more than extremes must be given, and yet it is not always possible to describe or reproduce a series in detail. Some form of abbreviation must be used.

A convenient method of summarizing data is what may be called the "cumulative- or moving-range." If the time series of Chapter VIII is used, some such statement as the following may be prepared. Of course, the amount of detail can be varied to suit the needs of the problem.

TABLE A

TABLE ILLUSTRATING THE CUMULATIVE- OR MOVING-RANGE METHOD OF SHOWING DISPERSION IN HISTORICAL SERIES

Years	Importations	
	Amounts in (000's) lbs.	Per cent
1895 to 1913	1,421,152	100.0
1895 to 1900	326,797	23.0
1895 to 1905	656,368	46.2
1895 to 1910	1,075,752	75.7

The data may be put in this manner :

1895 to 1913	1,421,152	100.0
1910 to 1913	431,437	30.4
1905 to 1913	825,293	58.1
1900 to 1913	1,161,753	81.7

Applying the same method to the frequency series in Table M, Chapter VIII, the arrangement will be somewhat as follows :

TABLE B

TABLE ILLUSTRATING THE CUMULATIVE- OR MOVING-RANGE METHOD OF SHOWING DISPERSION IN FREQUENCY SERIES

AMOUNTS	FREQUENCIES	
	Amounts	Per cents
As much as \$5 but less than \$15.00	434	100.0
As much as \$5 but less than \$ 8.00	121	27.9
As much as \$5 but less than \$11.00	374	86.2
As much as \$5 but less than \$14.00	433	99.8
Or in this manner		
Less than \$15 but more than \$ 4.99	434	100.0
Less than \$15 but more than \$13.99	1	.2
Less than \$15 but more than \$10.99	60	13.8
Less than \$15 but more than \$ 7.99	313	72.1

This method consists of establishing a series of cumulations, the extent of the groups being successively widened. Grouping may be begun from either end and carried forward step by step. The thing that is striven for is a summary but one which characterizes the complete distribution. The method lends itself to arithmetic but not to diagrammatic or graphic presentation.

The range may be reduced to a relative basis, and the data relieved of the particular unit in which expressed, — that is, a coefficient may be established — by comparing the difference of the extremes with their sum. In the time series used above, this method gives a dispersion of $\frac{121{,}852{,}000 \text{ lbs.} - 46{,}631{,}000 \text{ lbs.}}{121{,}852{,}000 \text{ lbs.} + 46{,}631{,}000 \text{ lbs.}}$, or .45. The same method in the frequency series gives a coefficient of $\frac{\$15 - \$5}{\$15 + \$5}$, or .50. That is, the dispersion in the two cases when all deviations are considered is approximately equal.

TABLE C

TABLE SHOWING THE DECILS OF RELATIVE WHOLESALE PRICES IN THE UNITED STATES, BY YEARS — 1890–1910
(Taken from Mitchell, W. C., *Business Cycles*, p. 112)

Years	Lowest Relative Price	1st Decil	2nd Decil	3rd Decil	4th Decil	5th Decil (Median)	6th Decil	7th Decil	8th Decil	9th Decil	Highest Relative Price
1890	86	97	101	105	108	112	116	119	126	133	160
1891	74	99	101	105	109	111	113	116	122	132	158
1892	61	92	99	101	104	107	108	111	114	118	141
1893	70	90	96	100	102	104	106	109	111	119	158
1894	46	79	85	91	94	96	99	101	103	111	129
1895	53	79	86	88	91	94	95	98	100	105	149
1896	39	71	79	85	88	90	92	95	98	100	142
1897	56	71	78	85	88	91	93	95	98	102	128
1898	48	77	84	87	91	94	96	99	101	108	155
1899	46	86	89	94	97	100	103	108	112	129	149
1900	59	90	98	102	106	109	113	118	123	136	192
1901	49	90	97	101	104	107	111	115	120	133	222
1902	45	91	98	102	107	110	114	119	134	145	194
1903	43	90	98	104	108	111	114	121	129	143	192
1904	60	91	98	103	106	112	117	120	130	143	197
1905	59	85	97	104	110	114	120	126	131	149	238
1906	62	89	100	108	114	119	124	131	137	159	279
1907	42	95	104	112	121	129	132	139	147	171	304
1908	45	89	102	107	113	119	124	130	139	156	228
1909	48	89	102	111	117	121	127	135	146	172	243
1910	48	86	103	112	118	124	132	144	154	187	363

(2) The "Decil" Method (Graphic) for Time Series

Professor Wesley C. Mitchell has employed the "decil" method of showing dispersion in connection with price index numbers. It consists of plotting for successive periods the extremes as well as the nine decils of price changes. For

each period — year in the case chosen — the relative price changes for each commodity from year to year are arranged by years in an ascending order and the decils computed in the regular manner.[1] The fifth decil is, of course, the median. The distribution gives an excellent measure of scatteration or dispersion. The preceding table and the following Plate 22 show the manner in which the process is applied.

Commenting on this table, Mitchell says:

"In 1909, for example, one commodity had a relative price as low as 48, and another had a relative price as high as 243. Thus the arithmetic mean for that year, 121, represents relative prices which are scattered over a range of almost 200 points. But three-fifths of the 145 commodities had relative prices falling within a much narrower range — 44 points, the difference between the second and eighth decils — and one-fifth fell within limits of ten points — the difference between the fourth and sixth decils."[2]

By not being content to use a single expression such as the arithmetic mean, the median, or the mode, it is possible, by choosing decils, to show graphically over a period, and at each period included, the degree to which data are compressed or closely grouped around or are scattered away from their norm or central tendency.[3]

Professor Mitchell's relative price data constitute series of frequency distributions in which nothing more detailed is given than the decils and ranges. The merits of the method

[1] The formulæ for computing decils are, respectively, for 1st, 2d, 7th, — $\frac{n+1}{10}$; $\frac{2(n+1)}{10}$; $\frac{7(n+1)}{10}$. In all cases by n is meant the number of items.

[2] Mitchell, Wesley C., *Business Cycles*, p. 109, University of California Studies.

[3] A slight variation of this method already described in another connection has more recently been applied by Professor Mitchell to price changes in *Wholesale Prices in the United States*. Either method has great possibilities for the use in question. *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 173, July, 1915, Washington, D. C.

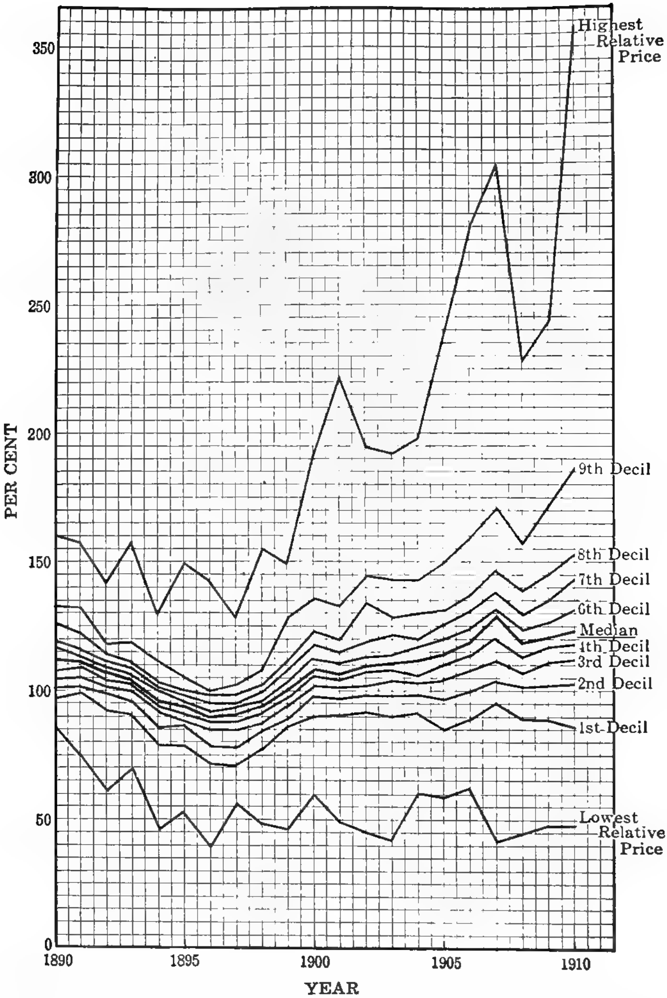


PLATE 22

Curves showing, by the Range and the Decil Methods, the Dispersion of the Fluctuations in Relative Wholesale Prices of 145 Commodities, 1890–1910.

consist in having data placed side by side — decil by decil — according to chronology, thus giving a continuous and detailed view of the spread or scatter. Not only may distribution be studied at a single period but also for all periods.

Whether the graphic or simply the arithmetic method of showing dispersion is used, comparison is made by noting differences. The facts in Table C above might be emphasized by showing successively the differences between the decils for the various years, and a summary, in the form of an average of some type, for the whole period. Other methods may be devised to make them emphatic.

(3) The Average Deviation

In order to compute deviations, obviously a standard must be adopted from which measurements are made. The mode, the median, or the arithmetic mean serves this purpose. If the arithmetic mean is used, and signs are considered, the differences are equal to zero. This follows as a matter of course from the nature of such an average. If, however, signs are disregarded, the aggregate deviations are larger when taken from the arithmetic mean than when taken from any other average, for the reason that this average is affected by both the size of the items and the frequencies. In the case of the median, however, they are a minimum — that is, are smaller than when calculated from any other average. Only the frequencies and the size of the items at or near the center of a distribution affect this measure. By the use of an analogy, Bowley has shown that the sum of the deviations is a minimum when calculated from the median.

"That the sum of the first powers is a minimum can be readily demonstrated, most easily by an analogy. Suppose that it is

required to run from a telephone exchange separate wires to everyone of n places in a straight line, where should the exchange be placed, so as to use the least total amount of wire? At the median position. For if you move from the median position to the right or to the left, you will find immediately that you are adding more wire than you are subtracting. Supposing there are 20 stations, and you have a position between the 10th and 11th; if you move to a position between the 11th and 12th, you have to increase your distance from 10 stations and diminish it from 9, in every case by the same length of the wire. The wires correspond to the deviations; and the sum of lengths of the wires is the sum of the lengths of the deviations. Consideration of this illustration will show that the sum of the deviations is a minimum when they are measured from the median, but that the median is not quite determinate, for if there are an even number of stations, the sums of the deviations measured from all points between the two central stations are the same.[1]

Mathematical consistency seems to demand that the median be used. On the other hand, the *average* deviation requires that the total be averaged, that is, divided by the number of items, and logical consistency seems to demand that they be computed from the mean.[2] In the examples following, the arithmetic mean is used both as standard and as divisor.[3]

The average deviation is an average. It is not different in this respect from the average of the original data. It does not represent a series of deviations in detail, but only attempts to record a type. When they are uniform and small, it does this satisfactorily. When they are large and different, it fails here as it does in the original case. Moreover, it is impossible to determine from the average alone which condition maintains. To do so requires that they be

[1] Bowley, A. L., *Measurement of Groups and Series*, p. 30.

[2] In symmetrical distributions and those only moderately asymmetrical the difference in the aggregate in the two cases would be small.

[3] Defense might be found for taking the *median* deviation when deviations are calculated from the median.

arranged into frequency groups or that the method of cumulative- or moving-ranges be used. When this is necessary must be determined by the data and the purposes for which they are used.

In the following examples the method of computing the average deviation is fully illustrated.

a. The Average Deviation in Historical Series

The following table gives the quantity of tin plates imported into the United States, 1906–1915, inclusive, in millions of pounds.

TABLE D

TABLE SHOWING THE QUANTITY OF IMPORTED TIN PLATES INTO THE UNITED STATES, 1906–1915, INCLUSIVE,[1] IN MLLIONS OF POUNDS

YEARS	AMOUNT	FREQUENCIES	DEVIATIONS		
			From average, 86.6		Total (signs ignored)
			–	+	
Total	86.6 (av.)	10	251.4	251.4	502.8
1906	121	1		34.4	251.4
1907	143	1		56.4	
1908	141	1		54.4	
1909	117	1		30.4	
1910	154	1		67.4	
1911	95	1		8.4	
1912	7	1	79.6		251.4
1913	28	1	58.6		
1914	49	1	37.6		
1915	11	1	75.6		

[1] *Statistical Abstract of the United States*, 1915, p. 498.

502.8
50.28

By disregarding signs and combining the deviations the total is 502.8. The average is therefore 502.8 ÷ 10 = 50.28. That is, the average difference of the various amounts imported from the average imported is 50.28 million pounds. The average itself is 86.6 million pounds. In one year the average is exceeded by 67.4 million pounds, while in another year the average imported exceeds the amount brought in in that year by 79.6 million pounds. The excess of the first is 78 per cent, and the deficit of the second 92 per cent, of the average. The average difference is 58 per cent of the average imported.

These differences might be illustrated in the following manner:

TABLE E

TABLE SHOWING IN CLASSIFIED FORM THE DIFFERENCES FROM THE AVERAGE IMPORTATIONS OF TIN PLATES INTO THE UNITED STATES

(Based on Table D)

DIFFERENCES FROM THE AVERAGE IMPORTATIONS (IN MILLION POUNDS)	YEARS IN WHICH THE CORRESPONDING DIFFERENCES WERE FOUND		
	Total	−	+
Total 86.6 (average)	10	4	6
Less than 15.0	1	—	1
15 but less than 30.0	—	—	—
30 but less than 45.0	3	1	2
45 but less than 60.0	3	1	2
60 but less than 75.0	1	—	1
75 but less than 90.0	2	2	—

Summarizing this table, it is shown that the positive and the negative differences from the average range from 90

to below 15 million pounds, six of the frequencies, when the deviations are taken positively, being between 30 and 60 million. The median difference when interpolated for is 55.4.

The average deviation may also be computed from an assumed average. The following table using the above data illustrates the method.

TABLE F

TABLE SHOWING THE METHOD OF COMPUTING THE AVERAGE DEVIATION WHEN AN ASSUMED AVERAGE IS USED

(Data same as in Table D)

YEAR	AMOUNT	FREQUENCIES		DEVIATIONS FROM ASSUMED AVERAGE — 90		TOTAL (SIGNS IGNORED)
				−	+	
Total	866		10	265	231	496
1906	121	1	6		31	231
1907	143	1			53	
1908	141	1			51	
1909	117	1			27	
1910	154	1			64	
1911	95	1			5	
1912	7	1	4	83		265
1913	28	1		62		
1914	49	1		41		
1915	11	1		79		

The total error in deviations is 34 — the difference between 265 and 231. Had they been computed from the true average the difference would have been zero. The average error is, therefore, 34 ÷ 10 or 3.4. Six of the frequencies

are too small — they were computed from 90 in place of 86.6 — and four of them are too large for the same reason. Therefore $(6 \times 3.4) - (4 \times 3.4)$, or 6.8, must be added to the combined frequencies, 496, to make up for the error. This gives 502.8 as the correct sum of the deviations when taken positively. The average deviation is therefore $502.8 \div 10$, or 50.28, as in the first method above.

There is no presumption of a normal or ideal arrangement in a time series. The average deviation, therefore, loses some of the significance associated with it in the treatment of natural phenomena. In the case of economic statistics it may be highly artificial. By its very nature the differences are important not only because of their size but also because of their distance from the center of gravity. In the example above the deviation of 8.4 is as important in the divisor as is that of 79.6. Each constitutes one of the ten differences. Of course, the median and the mode are differently affected.[1]

b. The Average Deviation in Frequency Series

In the discussion of the average deviation for frequency series there is no necessity of restating the essential differences between those that are discrete and continuous in type. What has already been said in this respect applies here. The present task is to comprehend its meaning and see its application to economic and business facts when they are grouped in frequency series.

Various types of frequency distributions are shown on Plate 23. Even on casual inspection, it is evident that it is futile to attempt to summarize them by a single expression such as an average. The averages may be similar, but the distributions about them widely different. It is the latter which

[1] See what is said relative to this point in Chapter VIII, *supra*.

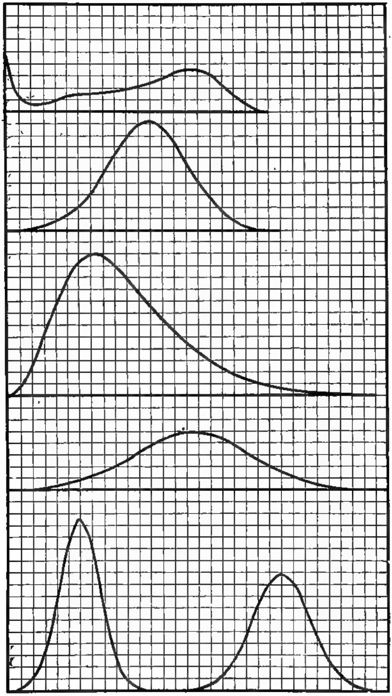


PLATE 23

Types of Frequency Distributions

are now being considered. Taking a somewhat different series, the application is seen in the following examples:

TABLE G

TABLE SHOWING THE METHOD OF COMPUTING THE AVERAGE DEVIATION IN A SIMPLE FREQUENCY DISTRIBUTION

AMOUNT	FREQUENCIES	DEVIATIONS				TOTAL (signs ignored)
		From True Average, $4.23		Multiplied by the Frequencies		
		−	+	−	+	
Total	37			$25.33[1]	$25.32[1]	$50.65
$2.00	4	$2.23		8.92		8.92
4.00	3	.23		.69		.69
3.00	9	1.23		11.07		11.07
6.00	5		$1.77		8.85	8.85
3.00	2	1.23		2.46		2.46
8.00	3		3.77		11.31	11.31
5.00	6		.77		4.62	4.62
3.50	3	.73		2.19		2.19
4.50	2		.27		.54	.54

Ignoring signs, the differences amount to $50.65. The average difference is, therefore, $50.65 ÷ 37, or $1.37. That is, the average difference from the arithmetic average is 32 per cent of the average, and varies, when weighted according to its importance, from the smallest positive difference of $.54 to the largest negative difference of $11.07.

The manner in which the average deviation is computed

[1] This negligible difference is due to taking the average as $4.23 rather than as $4.22 +.

for a series when the frequencies apply to groups is to assume for each group a uniform distribution, or what is the same thing, to assume that they are concentrated at the middle points, and proceed as in the case above. The following table, using a different set of data, is illustrative.

TABLE H

TABLE SHOWING THE METHOD OF COMPUTING THE AVERAGE DEVIATION FROM A GROUP-FREQUENCY SERIES

Amounts	Frequencies	Deviations: From the Average, $9.04 −	Deviations: From the Average, $9.04 +	Deviations: Product of Deviations and Frequencies −	Deviations: Product of Deviations and Frequencies +	Total Deviations (signs ignored)
Total . . .	434			$305.48 [1]	$305.12 [1]	$610.60
$5.00 to $5.99	15	$3.54		53.10		53.10
6.00 to 6.99	40	2.54		101.60		101.60
7.00 to 7.99	66	1.54		101.64		101.64
8.00 to 8.99	91	.54		49.14		49.14
9.00 to 9.99	113		$.46		51.98	51.98
10.00 to 10.99	49		1.46		71.54	71.54
11.00 to 11.99	30		2.46		73.80	73.80
12.00 to 12.99	27		3.46		93.42	93.42
13.00 to 13.99	2		4.46		8.92	8.92
14.00 to 14.99	1		5.46		5.46	5.46

The sum of the deviations is $610.60, and the average deviation $1.41. In this case, because of the concentration in the group $9.00 to $9.99, the average deviation is not much

[1] This negligible difference is due to taking the average to be $9.04 rather than $9.039+.

larger than the extent of this group, and is only 16 per cent of the average from which the deviations are computed. The figure unmistakably shows concentration, but it does not localize it.

If the differences are calculated from an assumed average, it is necessary to make correction for the difference between the guessed and the true average. The manner in which this is done in frequency series is shown in the following table:

TABLE I

Table Showing the Method of Computing the Average Deviation in a Group-Frequency Series when an Assumed Average is Used

Amounts	Frequencies		Deviations				
			From Assumed Average, $9.50		Product of Deviations and Frequencies		Total Deviations (signs ignored)
			−	+	−	+	
Total . . .		434			$403.00	$203.00	$606.00
$5.00 to $5.99	15	212	$4.00		60.00		60.00
6.00 to 6.99	40		3.00		120.00		120.00
7.00 to 7.99	66		2.00		132.00		132.00
8.00 to 8.99	91		1.00		91.00		91.00
9.00 to 9.99	113	222					
10.00 to 10.99	49			$1.00		49.00	49.00
11.00 to 11.99	30			2.00		60.00	60.00
12.00 to 12.99	27			3.00		81.00	81.00
13.00 to 13.99	2			4.00		8.00	8.00
14.00 to 14.99	1			5.00		5.00	5.00

The total error in deviations is \$200.00 — the difference between \$403.00 and \$203.00. The average error is, therefore, \$200.00 ÷ 434 or \$.461. But 212 of the frequencies are too large since they were computed from \$9.50 instead of \$9.04; and 222 of them are too small for the same reason. Therefore, the difference between 212 × \$.461 and 222 × \$.461 must be added to the total frequencies — \$606.00 — in order to get the correct total. \$606.00 − (212 × \$.461) + (222 × \$.461) = \$610.60, and this divided by the number of instances, 434, equals \$1.41, the correct average deviation.

TABLE J

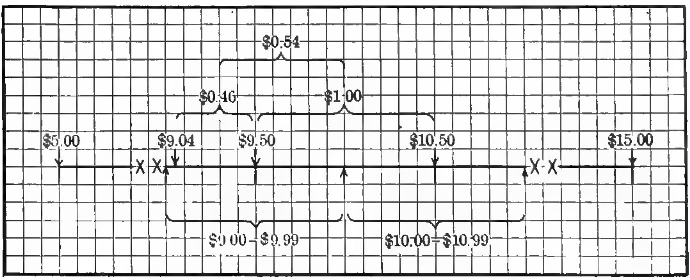
TABLE SHOWING THE METHOD OF COMPUTING THE AVERAGE DEVIATION IN A GROUP-FREQUENCY SERIES FROM AN ASSUMED AVERAGE BY THE "STEP-DEVIATION" METHOD

AMOUNTS	FREQUENCIES		DEVIATIONS IN "STEPS" — From Assumed Average, \$10.50		Product of Deviations and Frequencies		Total (signs ignored)
			−	+	−	+	
Total . .		434			728	94	822
\$5.00 to \$5.99	15	212	5		75		75
6.00 to 6.99	40		4		160		160
7.00 to 7.99	66		3		198		198
8.00 to 8.99	91		2		182		182
9.00 to 9.99	113	113	1		113		113
10.00 to 10.99	49	109					
11.00 to 11.99	30			1		30	30
12.00 to 12.99	27			2		54	54
13.00 to 13.99	2			3		6	6
14.00 to 14.99	1			4		4	4

The so-called "step-deviation" method, used in Chapter VIII for computing the arithmetic mean, may be used in connection with the average deviation. Moreover, a consideration to be kept in mind when the method employed in Table G is used, may be explained. Suppose an average of $10.50 is assumed and the average deviation is calculated for the above series by the "step" method. The preceding table shows the result.

The total error in step-deviations is 634; the difference between 728 and 94. The average step-deviation error is, therefore, 634 ÷ 434 or 1.46. The steps are all of $1.00 width, so that the average step-deviation error, in terms of the unit of measurement, is $1.00 × 1.46 or $1.46. But the combined deviations, 822, are computed from $10.50 instead of $9.04, the true average. Some of them are too small and some are too large. Which are affected and how much? The frequencies above $8.50 are each too large by $1.46 on the average. Those at $10.50 and below are each too small by the same amount. Those at $9.50, 113, are each too large by $1.00 if $10.50 is used. But, $9.04 instead of $9.50 is the average. Therefore, each of 113 is too large by the difference between $1.00 and $.46, which is $.54.[1]

[1] The reason for an overlapping is shown by diagram below:



The total deviations properly corrected are $822 - (212 \times \$1.46) + (109 \times \$1.46) - (113 \times \$.54)$ which equals \$610.6. The average deviation is, therefore, $\$610.6 \div 434$ or \$1.41.

This seems a roundabout method of reaching a simple result. It is true only when the guessed average falls outside of the limits of the group which contains the true average. If it falls within this group, the method is simple and possesses merits for some uses.

So much for the method of computing the average deviation in both time and frequency series. Just a word of recapitulation. The average deviation is an *average*. It does not necessarily reflect the peculiarities of deviations any more than the arithmetic mean does of data out of which it is computed originally, except for the fact that the respective variations from the average deviation are usually not as large as are the variations of the original data from their average. If it is large it shows relative dispersion; if it is small it shows relative concentration. The exceptions are weighted in this case in the same way that they are in any arithmetic mean. If the median or modal deviations are used, then they exercise less weight. If the cumulative-range method is used, they are thrown into prominence in detail. The need for a single summarizing expression in many economic and business fields is by no means so pressing, nor is its application so clear, as it is in the field of natural science.

Average deviations may be reduced to a comparable basis by dividing them by the averages from which they are computed. By so doing data are rid of the units in which expressed, and comparisons made possible. That is, coefficients are established. The coefficient in the case of the frequency distribution used as an example, since the differ-

ences were calculated from the arithmetic mean, is $\frac{\$1.41}{\$9.04}$ or .156.[1]

(4) The Standard Deviation

The standard is a modification of the average deviation. It is computed by measuring the respective deviations from the arithmetic average, by squaring these, thus getting rid of the minus signs, by averaging the total, and finally by extracting the square root. In the formula n refers to the number of instances — frequencies; d^2, to the deviations squared; and Σ, to the sum of the products of the frequencies and the squares. It is usually indicated by small sigma, σ, or by S. D., and by the formula $\sqrt{\frac{\Sigma(d^2)}{n}}$. Squaring gives weight to extremes — those deviations far removed from the average. This is not fully compensated for in the subsequent root extraction. In frequency distributions which follow the normal law of error, or which are moderately asymmetrical, instances far removed from the average are relatively few, so that the products of the squares and the frequencies at these points are due more to the squaring than to the multiplication. Near the average, however, frequencies are relatively numerous and the products affected by the concentration. In averaging the squares of the deviations, the frequencies, as such, exert equal weight, since the total is simply divided by the sum of the frequencies.

In time or historical series the case is somewhat different. There is no multiplication of deviations by frequencies, since

[1] On the graphic method of indicating absolute and relative dispersion, see Clark, Earle, "The Horizontal Zero in Frequency Diagrams," in *Quarterly Publications of the American Statistical Association*, June, 1917, pp. 662–669.

each item appears but once. The squaring alone is effective. Of course, distance from the average is still important, but this is neither accentuated nor minimized by the distribution of frequencies. Just as the sum of the deviations is a minimum, — that is, least, — when calculated from the median, so the sum of the squares of the deviations is a minimum when calculated from the arithmetic mean. This follows from the principle that the nearest approach to the mathematically correct measure or observation in a series is the arithmetic mean, and that errors in observation are distributed about this center according to the rule of squares.[1]

For many economic and business purposes interest lies chiefly in the thing that is characteristic. Legislation is not generally enacted for the few, but rather for the many. Business policies are most frequently mapped out and changed in light of that which seems to be characteristic. Sometimes, however, it is the exception which is suggestive, or which calls attention to the need for change. For instance, an exceptionally large sale — that far removed from the characteristic performance — may suggest possibilities in management and deserve to be emphasized both because of its stimulating effect on future performances on the part of salesmen, and because of its suggestive power to the management as to the need of reorganization of the selling force. Wide dispersion of employees' earnings in piece-work establishments may suggest to a keen business management the possibilities of a redistribution of labor service according to capacity and proved ability. The losses resulting from a haphazard use of labor force, when measured in terms of discontent, turnover of labor, etc., may well make it advisable to assign more importance to the exception than that which would follow from its mere numerical significance. The in-

[1] See Yule, G. Udny, *Introduction to the Theory of Statistics*, pp. 134–135.

equalities of wealth distribution carry with them a significance far greater than that indicated by amounts alone.

So long as it is desired to give moderate weight to large differences, the average deviation may be used. When interest shifts to that which is exceptional, means of throwing it into light are needed. Of course, in statistics of economics and business there is generally no presumption of normal distribution as there is in statistics of natural phenomena. Interest in deviations from type in the two cases is of a different kind. Respecting the latter, deviations are important as showing non-conformity to an abstract standard; respecting the former, as means of calling attention, for instance, to useless waste, to unnecessary sources of industrial disorder, etc. Approach in the two cases may be different, but the means of measuring the concentration or dispersion be the same. To cite an average alone is frequently inadequate in economics, even for general purposes. But to use both an average and the standard deviation gives a rather definite idea of distribution about this figure. The latter serves more accurately to define the average. Moreover, average and standard deviations bear a more or less definite relation to each other in distributions which approach the normal law. As Yule says,

> "It is a useful empirical rule for the student to remember that for symmetrical or only moderately asymmetrical distributions, approaching the ideal forms . . ., the mean deviation is usually very nearly four-fifths of the standard deviation."[1]

Again, the standard deviation is more or less definitely fixed. Respecting this Yule says:

> "It is a useful empirical rule to remember that a range of six times the standard deviation usually includes 99 per cent

[1] Yule, G. Udny, *Introduction to the Theory of Statistics*, p. 146.

or more of all the observations in the case of distributions of the symmetrical or moderately asymmetrical type."[1]

How nearly this is true for the frequency distributions chosen for example is evident on inspection.

a. The Standard Deviation in Historical or Time Series

Using the time series of Table D, the standard deviation is computed as follows, when the direct method is used:

TABLE K

TABLE SHOWING THE METHOD OF COMPUTING THE STANDARD DEVIATION FOR HISTORICAL SERIES USING THE DIRECT METHOD

(Data same as in Table D)

Years	Amount	Frequencies	Deviations From Average, 86.6		Deviations Squared	Deviations Squared, Multiplied by Frequencies
			−	+		
Total	86.6(av.)	10				29,760.40
1906	121	1		34.4	1,183.36	1,183.36
1907	143	1		56.4	3,180.96	3,180.96
1908	141	1		54.4	2,959.36	2,959.36
1909	117	1		30.4	924.16	924.16
1910	154	1		67.4	4,542.76	4,542.76
1911	95	1		8.4	70.56	70.56
1912	7	1	79.6		6,336.16	6,336.16
1913	28	1	58.6		3,433.96	3,433.96
1914	49	1	37.6		1,413.76	1,413.76
1915	11	1	75.6		5,715.36	5,715.36

[1] *Ibid.*, p. 140.

The deviations squared and totaled amount to 29,760.40. The standard deviation is, therefore, $\sqrt{\frac{29{,}760.40}{10}}$ or $\sqrt{2{,}976.04}$ or 54.5. The average deviation, 50.28, is 92.3 per cent of this amount.

TABLE L

TABLE SHOWING THE METHOD OF COMPUTING THE STANDARD DEVIATION FOR HISTORICAL SERIES USING THE DIRECT METHOD BUT AN ASSUMED AVERAGE

(Data same as in Table D)

YEARS	AMOUNT	FREQUENCIES	DEVIATIONS			
			From Assumed Av., 90.0		Squared	Squared, Multiplied by Frequencies
			−	+		
Total	86.6(av.)	10				29,876
1906	121	1		31	961	961
1907	143	1		53	2,809	2,809
1908	141	1		51	2,601	2,601
1909	117	1		27	729	729
1910	154	1		64	4,096	4,096
1911	95	1		5	25	25
1912	7	1	83		6,889	6,889
1913	28	1	62		3,844	3,844
1914	49	1	41		1,681	1,681
1915	11	1	79		6,241	6,241

In this example, the deviations are taken from the assumed average, 90.0, instead of the true average, 86.6. The average error in deviations is, therefore, 3.4. This must be squared and multiplied by the number of frequencies and then sub-

tracted from 29,876 in order to get the correct deviations squared. The square of 3.4 is 11.56, and when multiplied by 10 — the number of frequencies — is 115.6. The difference between this and 29,876 is 29,760.4. The square root of this amount is 54.5 and is the standard deviation. The problem is somewhat simplified by taking the deviations from an assumed average since the numbers to be squared are even. Of course, in actual work it is unnecessary to go through the form of multiplying by the frequencies when they are all unity. It was done here in order that all the steps might be followed.

TABLE M

TABLE SHOWING THE METHOD OF COMPUTING THE STANDARD DEVIATION FOR FREQUENCY SERIES BY USING THE SHORT-CUT METHOD AND AN ASSUMED AVERAGE

(Data same as in Table I)

Amounts	Frequencies	Deviations			
		From Assumed Av., $9.50		Squared	Squared, Multiplied by Frequencies
		−	+		
Total	434				$1,424.00
$ 5.00 to $ 5.99	15	$4.00		$16.00	$240.00
6.00 to 6.99	40	3.00		9.00	360.00
7.00 to 7.99	66	2.00		4.00	264.00
8.00 to 8.99	91	1.00		1.00	91.00
9.00 to 9.99	113				
10.00 to 10.99	49		$1.00	1.00	49.00
11.00 to 11.99	30		2.00	4.00	120.00
12.00 to 12.99	27		3.00	9.00	243.00
13.00 to 13.99	2		4.00	16.00	32.00
14.00 to 14.99	1		5.00	25.00	25.00

b. The Standard Deviation in Frequency Series

The method of calculating the standard deviation is the same for frequency as for time series, but it may be helpful to carry through an example when the direct and the indirect methods are employed. Taking the data in Table I, and assuming the average to be \$9.50 — the true average being \$9.04 — the short-cut method is as shown in Table M, on the preceding page.

The sum of the squares of the deviations from the guessed or assumed average is \$1,424.00. But the average error is \$.461. The square of \$.461 is \$.212. This amount multiplied by the number of frequencies — 434 — gives \$92⁺, and this amount, when subtracted from \$1424, gives \$1332, as the correct deviations squared. But since it is the average deviation that is desired it is necessary to divide this number by 434. The result is \$3.07. The square root of \$3.07 is \$1.75 and is the standard deviation. The average deviation — \$1.41 — is 81 per cent of this amount.

The standard deviation of a series is somewhat larger than its average deviation. If the distribution is normal in the probability sense, the two measures of variability stand in the following relation:

$$\sigma \text{ or S.D.} = 1.2533 \text{ A.D., or conversely,}$$
$$\text{A.D.} = 0.7979\, \sigma \text{ or S.D.}$$

Applying this formula to the example used as an illustration, the relation between the average and the standard deviations is as 1 : 1.2413, or conversely 0.8056^{+} : 1. Inserting these quantities in the formulæ,

$$\sigma = 1.2413 \text{ A.D., or conversely,}$$
$$\text{A.D.} = 0.8056\, \sigma$$

That is, the distribution approaches very nearly the normal or probability curve.

If the same distribution and a guessed average are used and the deviations are taken in terms of "steps," the method is the same, except that it is necessary to convert the steps into terms of the unit employed by multiplying by the size of the group. In this case the step is \$1.00. If the widths of groups had been \$.50, for instance, the conversion would have been made by multiplying the number of steps by one half dollar.

If deviations from the actual average, as they appear in Table H, are used, the process is the same but the chance of error greater since they are larger and more difficult to square. Of course, in such case it is unnecessary to make correction for the error in deviations. They are correct by assumption.

In order to convert the standard deviation into a *coefficient* — that is, to relieve the data from the particular unit in which expressed and to make comparisons possible between two series in which absolute units are different — it is only necessary to divide by the arithmetic mean — the figure from which the deviations are computed. The *coefficient* of dispersion for this series based on S.D., is $\frac{\$1.75}{\$9.04}$ or .194.

(5) The Quartile Measure

The quartile measure of dispersion applies to that portion of a distribution contained between the first and third quartiles. The extremes below the first and beyond the third quartiles are ignored. It serves to characterize that portion which lies nearest the average or type. This measure, like the average and standard deviations, is an average,

but is not calculated from the differences from the median, mode, or arithmetic mean, but by taking one half of the range contained in the middle half of a distribution. The formula is $\frac{Q3 - Q1}{2}$, where $Q3$ and $Q1$ stand for the third and first quartiles, respectively. The third quartile lies above the median, the first one below. The distance they are apart, and the proportion of a complete distribution contained between them, are roughly indicated by this measure. If a distribution is symmetrical, this figure, when added to the lower or subtracted from the upper quartile, coincides with the median. If asymmetrical at all, of course, it will differ, the size indicating the place at which asymmetry appears. A rough measure of dispersion is found by comparing the range of the middle half with the complete range of a series, or the average range of the middle half with the average range of the first and last quarters. Other modifications of the quartile measure may be devised.

In symmetrical or moderately asymmetrical distributions the relation between the quartile and the standard deviation measures of dispersion is fairly constant and predictable. The first is generally about two thirds of the second, and nine times the first usually contains about 99 per cent of a total distribution.[1] How nearly the relationship maintains in the distribution chosen as an illustration is seen by the following: In Table M, Chapter VIII, the median, by interpolation, was fixed at $9.049. The first and third quartile positions, by the formula $\frac{n+1}{4}$, and $\frac{3(n+1)}{4}$, respectively, are the $108\frac{3}{4}$ and $326\frac{1}{4}$ men. The wages of these hypothetical individuals, when interpolated for, are $7.81 and $10.03, respectively. The quartile range is, therefore, $10.03 –

[1] Yule, *op. cit.*, p. 148.

\$7.81 or \$2.22. The average range is $\frac{\$2.22}{2}$, or \$1.11.[1] For the same series the average deviation is \$1.41, and the standard deviation \$1.75. The semi-quartile range, therefore, is equal to 79 per cent of the former and 63 per cent of the latter. The extreme range of \$10.00 — the difference between \$5.00 and \$15.00 — is almost exactly nine times the quartile measure, \$1.11.

As in the cases of other measures of dispersion the semi-quartile range may be reduced to a relative basis, or made a coefficient, by dividing through by a common denominator. In this case, the appropriate divisor is the sum of the quartiles. The fraction — $\frac{Q3 - Q1}{Q3 + Q1}$ — increases with the distance between the quartiles but always lies between 0 and 1. Size, therefore, is a test of relative dispersion. In the above example the coefficient is $\frac{\$10.03 - \$7.81}{\$10.03 + \$7.81}$, or .124. That is, the dispersion is relatively small. It is 79 per cent of the coefficient based on the average deviation and 64 per cent of the coefficient based on the standard deviation.

For many purposes a study of the semi-quartile range is sufficient. This may result from the nature of a distribution or from lack of interest in the extreme cases. However, to cite only this measure may prejudice a case for all purposes except those which are under discussion. In order to guard against misunderstanding and to give expression to all the peculiarities of a distribution, it is generally better to determine the average, the standard, and the quartile

[1] For discrete series, interpolation in units less than those in which data are measured is illogical and aims at too great accuracy. For most purposes the quartiles would be given with sufficient accuracy as \$7.80 and \$10.00.

deviations. A comparison of these gives an accurate picture of a distribution.

(6) The "Probable Error"

At this point it is necessary to introduce a different concept. Statistical studies are almost always made by using sample measurements. All prices cannot be included in computing an index number nor all rents determined when studying family budgets. Neither the time required for all operators within manufacturing industries to complete an operation, nor the time necessary for every operator in telephone industries to answer the telephone calls of all subscribers, can be determined in order to answer a specific inquiry. Sample measurements must be used and some method employed for testing the reliability of those taken. Averages per se will not suffice; their limitations in describing frequency distributions are clear. The most common measure of divergence from type is the standard deviation. But it is simply a measure for the samples taken. What is wanted is proof that the distribution in the samples taken indicates the distribution that would result if the whole "population"[1] were included. The probable error supplies this. On the supposition that if all the population were included a distribution would follow the normal curve of error, the probable error stands in a mathematical relation to the standard deviation in the same way that the radius of a circle does to the circumference. When this distribution does not maintain, of course, the relationship no longer holds.

For a probability distribution the probable error is approximately two thirds of the standard deviation, or more

[1] "Population" is a word used to indicate the complete group, samples of which are measured.

exactly P.E. = 0.6745 σ. It is a "pair of values lying one above and the other below the value determined. We can say that there is an even chance that the true value lies between these limits."[1]

Jevons has illustrated the concept as follows:

"Suppose, for instance, that five measurements of the height of a hill . . . have given the numbers of feet as 293, 301, 306, 307, 313; we want to know the probable error of the mean, namely 304. Now the difference between the mean and the above numbers, paying no regard to directions, are 11, 3, 2, 3, 9; their squares are 121, 9, 4, 9, 81, and the sum of the squares of the errors consequently 224. The number of observations being 5, we divide by 1 less, or 4, getting 56. This is the square of the mean error, and taking its square root we have 7.48 (say $7\frac{1}{2}$), the mean error of a single observation. Dividing by 2.236, the square root of 5, the number of observations, we find the mean error of the *mean* result to be 3.35, or say $3\frac{1}{3}$, and lastly multiplying by .6745, we arrive at the *probable error of the mean result*, which is found to be 2.259, or say $2\frac{1}{4}$. The meaning of this is that the probability is one half, or the odds are even that the true height of the mountain lies between $301\frac{3}{4}$ and $306\frac{1}{4}$. We have thus an exact measure of the degree of reliability of our mean result, which mean indicates the most likely point for the truth to fall upon."[2]

Whipple defines it as follows:

"The probable error, P. E., of a single measure is an amount of deviation both above and below M (or median or mode) that will

[1] Davenport, C. B., *Statistical Methods*, p. 14.
"The chances that the true value lies within

$\pm$ 2 E are	4.5	: 1
$\pm$ 3 E are	21	: 1
$\pm$ 4 E are	142	: 1
$\pm$ 5 E are	1,310	: 1
$\pm$ 6 E are	19,200	: 1
$\pm$ 7 E are	420,000	: 1
$\pm$ 8 E are	17,000,000	: 1

$\pm$ 9 E are about a billion to 1." *Ibid.*

[2] Jevons, W. Stanley: *The Principles of Science*, p. 388 (2d Edition).

include one-half of the individual measures; that is, it is a value such that the number of deviations that exceed it (in either direction from M) is the same as the number of deviations that fall short of it." [1]

Pearl, speaking of it in a different application, says:

"Suppose that we read that the mean length of the thorax of a thousand fiddler crabs is 30.14 ± .02 mm. Just what does this actually mean? Accepting the figures at their face value, or, put another way, assuming that the mathematical theory on which the probable error was calculated was the correct one, the figures mean something like this: If one were to take, quite at random, successive samples of 1000 each from the total population of fiddler crabs and determine the mean thoracic length from each sample, these means would all be different from each other by varying amounts. In other words, no single sample would give us the absolutely true value of the mean thoracic length of the fiddler crab population. The true value is in an absolute sense unknowable, because, for one reason, always we must come at the finding of it by way of random sampling, and sampling means variation. Now it is an observed fact of experience that the variations due to random sampling distribute themselves according to a definite law of mathematical probability. Knowing this law, it is clearly possible to state the mathematical probability for (or against) any particular deviation or variation occurring as the result of random sampling. Exactly this is what the probable error does. It says, in the particular case here considered, that it is an even chance, that a deviation or variation in the value of the mean as great or greater than .02 mm. above or below will occur as a result of random sampling. Or, put in another way, if we took successive samples of 1000 each from this crab population, it is an even bet that the value of the mean from any sample would fall between $30.14 + .02 = 30.16$, and $30.14 - .02 = 30.12$." [2]

The probable error, therefore, is a means of testing the reliability of samples provided that data approach the nor-

[1] Whipple, Guy M., *Manual of Mental and Physical Tests*, Part 1, p. 23.
[2] Pearl, Raymond, *Modes of Research in Genetics*, pp. 96-97.

mal probability distribution. The probable error of a given deviation is then indicated by one half of the distance between the upper and the lower quartiles, *i.e.* the quartile measure of deviation furnishes a measure of the likelihood that a deviation will fall within one half of the distance above or below the median.[1] Referring again to the distribution in Table M, Chapter VIII, the semi-quartile range was found to be \$1.11, and the standard deviation \$1.75.[2] Applying the formula, P. E. $= 0.6745\ \sigma$, in this case the P. E. should have been \$1.18 rather than \$1.11. The computed, therefore, is 94.1 per cent of the theoretical probable error.

The probable error may likewise be computed for the arithmetic mean of a number of measurements, the *means* of which vary. Suppose it is desired to measure the length of time in which a certain manufacturing process is completed, or in which a given task is done, as a basis for task setting. If a large number of trials are made for homogeneous groups of operators and averages of the periods taken for each group, these will vary.[3] The standard deviation of the averages and its probable error may be taken in the same way that they are computed for single variations. The formula for the probable error of the mean is

$$\pm\ 0.6745 \times \frac{\text{Standard Deviation of the means}}{\sqrt{\text{Number of variates}}} \quad \text{or}$$

$$\pm\ 0.6745 \times \frac{\text{S.D.}}{\sqrt{n}}.$$

[1] For a normal distribution arithmetic mean and median coincide.

[2] *Supra*, p. 406.

[3] See the interesting account of the results of a series of experiments involving the accuracy with which estimation is made by trained employees. Harris, J. Arthur: "Experimental Data on Errors of Judgment in the Estimation of the Number of Objects in Moderately Large Samples, with Special Reference to the Personal Equation." *The Psychological Review*,

The meaning of such a figure is indicated above in the quotation from Professor Pearl.

A few instances where the probable error may be applied in economic studies may be cited. Breeders of animals and plants find constant need of using it in studies of variation from type and in correlation.[1] Moreover, in the selection of men according to psychological and other tests,[2] in the grading of cotton and grains, in the setting of tasks, and the establishment of piece-rates of compensation on the basis of the "average" operator's performance, some measure of the reliability of the samples must be employed. Again, according to some [3] the only scientific method of establishing the pure premium for industrial accident insurance is to compare homogeneous conditions of risk exposure and to test the homogeneity by measures of dispersion. Conformity to the normal law is proof that conditions are homogeneous. Most comparisons, it is held, involve non-homogeneous conditions. The proper unit is not the "establishment," but similar risk conditions in many establishments or industries.

In studies of correlation the probable error always accompanies the coefficient as a test of reliability. This phase of the problem is discussed later.[4]

It must be remembered that the probable error is to be used only when distributions approach the normal probability form and where samples are relatively numerous.

Vol. XXII, No. 6, November, 1915, pp. 490–511. In this series of experiments there is a clear tendency for the estimates to be too high.

[1] Davenport, Eugene, *The Principles of Breeding*, *passim*, New York, 1907.

[2] Whipple, Guy M., *Manual of Mental and Physical Tests*, Baltimore, 1914.

[3] Cf. Fisher, Arne, *Proceedings of the Casualty, Actuarial, and Statistical Society of America*, Vol. II, Part III, No. 6, May, 1916.

[4] See Chapter XII, *infra*.

The standard deviation, however, as a measure of divergence from the norm is of general application. As Yule says, "In the case of small samples, the use of the probable error is consequently of doubtful value while the standard error (deviation) retains its significance as a measure of dispersion." [1] However, "On the whole, the use of the 'probable error' is of little advantage compared with the standard. . . ." [2]

III. Skewness

1. *Meaning of Skewness*

Measures and coefficients of dispersion, both in historical and frequency series, indicate absolutely or relatively the differences of the separate measures from a single one taken as a standard. They represent deviations from type, varying emphasis being given to the differences depending upon the particular measure used. The average deviation gives all differences their normal weight; the standard deviation accentuates those far removed from type, but still averages them. The quartile measure includes only those lying within the boundaries of the first and third quartile. As such, none of them reveal the distributions of the deviations. Differences from the type are not localized. The degree to which they cluster above or below the type is not shown. What measures of skewness do is to localize the degree to which distributions are pulled, distorted, or skewed from normality, *i.e.* from the symmetrical form which they take when mode, median, and arithmetic mean coincide. The differences between these in themselves indicate asymmetry, that is, a piling up or scattering of frequencies on one or the other side of the type. These may be expressed relatively,

[1] Yule, G. U., *Introduction to the Theory of Statistics*, p. 307.
[2] *Ibid.*

so as to admit of comparison, by being reduced to coefficients. Measures of dispersion which characterize the distribution on both sides of the type must be used as divisors, since what is desired is a relative expression of the localization of asymmetry. To divide by the units in which the measures are expressed would be simply to reduce the deviations to a relative basis.

Distributions generally are skewed to some degree. Rarely if ever, even among natural phenomena, is complete symmetry found.[1] This may be due to the unrepresentativeness of the samples, to imperfect measurements, or to other causes. Distributions may be scattered widely or closely grouped, but rarely are they uniformly grouped or distributed about a norm. Measures and coefficients of skewness localize deviations from symmetry; measures and coefficients of dispersion only reveal the amount of scatteration or cluster.

2. *Measures and Coefficients of Skewness*

The chief and currently used measure of skewness is the difference between the arithmetic mean and the mode. If the mean exceeds the mode — that is, is drawn away from the typical instance by the presence of extreme items — skewness is said to be positive. If it is less than the mode — that is, is drawn away from the typical instance because of extreme items — skewness is said to be negative. The mode is unaffected by extremes, either small or large, except at or near the center of a distribution; while the arithmetic mean is not only affected by the size of the items but also by the distance away from the center of gravity. The dif-

[1] Cf. Tolley, Howard R., "Frequency Curves of Climatic Phenomena," in *Monthly Weather Review*, United States Department of Agriculture, Vol. 44, November, 1916, pp. 634–642, 636.

ference between these items, therefore, serves as a measure of skewness. Extending the same principle, both of the measures may be compared with the median. It is useful to note the empirical rule that in distributions which are moderately asymmetrical, the median travels about two thirds of the distance from the mode toward the arithmetic mean. If this relationship — mode = mean − 3 (mean − median) — is exceeded, skewness is marked; if the reverse is true, then skewness is small. It is localized by the relative positions of the three measures. In markedly asymmetrical series, the mode may be indeterminate, or it may be misplaced by the use of this formula. When there is no mode or when a series is bi-modal, it is difficult, if not impossible, to measure skewness by this simple rule.

The measure of skewness based on the difference between the mode and arithmetic mean may be reduced to a coefficient by dividing by the standard or average deviation, the formula in the first case being $\frac{\text{mean} - \text{mode}}{\text{S.D.}}$. The former is the more common divisor. If the mean is on the lower side of the mode, when the statistics are plotted in a diagram, this function is negative. If on the upper side, it is positive.

Taking the frequency series in Table H the arithmetic mean is \$9.04, the median, by interpolation, \$9.05, and the mode, by inspection, \$9.50. Skewness is, therefore, negative on the basis of the measure, mean − mode or \$9.04 − \$9.50, the coefficient being $\frac{-\$.46}{\$1.75}$ or − .26. Based on the average deviation, the coefficient is $\frac{\$.46}{\$1.41}$ or − .33.

As in the case of dispersion, measures and coefficients of skewness may be restricted to that portion of a distribution

falling between the first and third quartiles. Dispersion is then measured by the difference between these quarter measures divided by their sum. Skewness is localized by subtracting from the sum of the quartiles twice the median, and the coefficient, based on this measure, secured by dividing by the difference between the quartiles. In the example above, the first and third quartiles respectively were found to be \$7.81 and \$10.03. The median was placed at \$9.05. By the formula skewness is indicated by $\$7.81 + \$10.03 - (2 \times \$9.05)$ or $-.26$, and is negative. The coefficient is $\frac{-\$.26}{\$2.22}$ or $-.12$. That is, skewness is negative for the center one half of the distribution and is something less than one half of what it is for the complete series.

Measures and coefficients of both dispersion and skewness should be in everyday use in statistical work. For two or more series arithmetic means may be identical, but dispersion widely different; dispersion may be identical, but skewness different. These facts are important. A comparison of sales, wages, interest rates, stock and bond prices, by means of such measures could not fail to throw new light on the everyday problems of business.

Without carrying through the arithmetical steps in the computation of these factors for a typical problem (see Plates 24 and 25), since this would involve unnecessary repetition of the methods already given, their significance may be made real by using comparable wage data for a single occupation in eighteen identical establishments, reported by the United States Bureau of Labor Statistics. The following table gives the classified wage data and the summaries which have been computed from them:

TABLE N

TABLE SHOWING CLASSIFIED WAGE-RATES[1] OF FEMALE MENDERS IN EIGHTEEN IDENTICAL WOOLEN AND WORSTED MANUFACTURING ESTABLISHMENTS, BY YEARS, TOGETHER WITH CERTAIN MEASURES OF DISPERSION[2] AND SKEWNESS[2]

WAGE GROUPS — CENTS PER HOUR	CLASSIFIED WAGE-RATES OF FEMALE MENDERS, BY YEARS			
	1907	1908	1909	1910
Total	403	341	583	498
6 to 8	—	3	3	1
[3] 8 to 9	2	8	44	14
[3] 9 to 10	27	22	91	44
10 to 12	68	71	117	125
12 to 14	119	61	82	81
14 to 16	81	57	86	58
16 to 18	37	39	49	30
18 to 20	34	35	42	82
[4] 20 to 25	31	35	58	43
25 to 30	4	10	11	16
[5] 30 to 40				4
[6] 40 and over				
Arithmetic Mean	14.56¢	15.01¢	13.96¢	14.97¢
Mode (by interpolation)	13.08¢	(7)	10.95¢	(7)
First Quartile	12.07¢	11.48¢	10.14¢	11.05¢
Median (2d Quartile)	13.76¢	14.22¢	12.09¢	13.62¢
Third Quartile	16.32¢	17.77¢	16.61¢	18.52¢
Dispersion:				
Average Deviation	2.86¢	3.54¢	3.75¢	4.00¢
Standard Deviation	3.67¢	4.47¢	4.58¢	4.96¢
Coefficient on A. D.	.196	.236	.269	.287
Coefficient on S. D.	.252	.298	.328	.331
Skewness:				
Mode — Arithmetic Mean	+ 1.48¢	(7)	+ 3.01¢	(7)
Quartile Measure	+ .87¢	+ .81¢	+ 2.57¢	+ 2.33¢
Coefficient on S. D.	+ .40	(7)	+ .66	(7)
Coefficient on Quartile	+ .21	+ .13	+ .40	+ .31

[1] *Bulletin of the United States Bureau of Labor Statistics*, Whole Number 190, May, 1916, p. 139.

[2] Computed.

[3] Notice size of group.

[4] Notice size of group.

[5] Notice size of group.

[6] Notice residuum.

[7] Indeterminate.

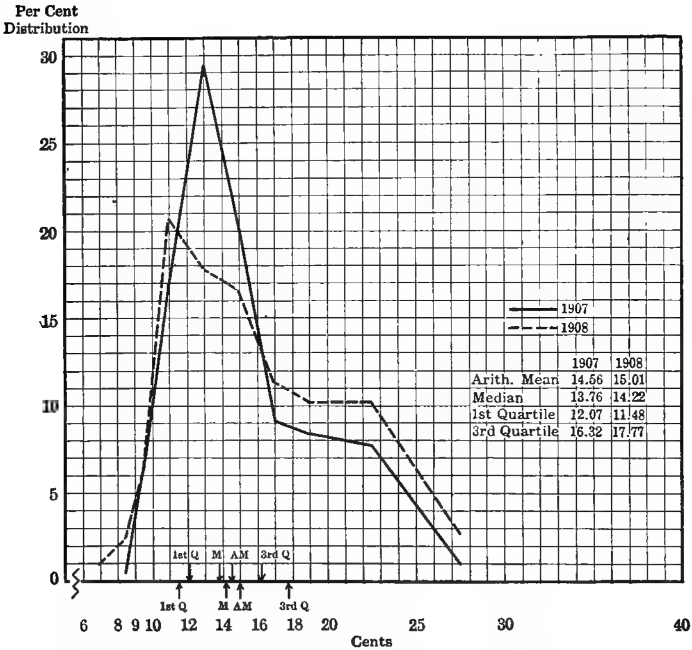


PLATE 24

Curves Showing, by Years, Classified Wage-rates of Female Menders in Woolen and Worsted Establishments.

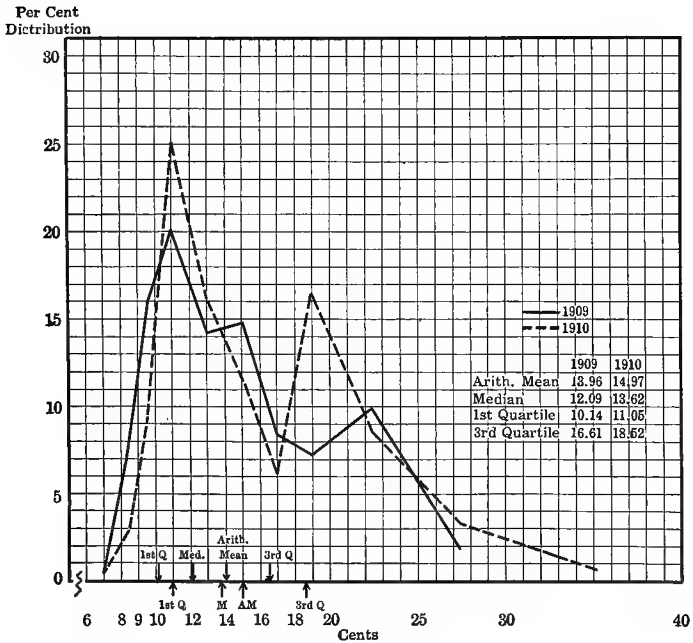


PLATE 25

Curves Showing, by Years, Classified Wage-rates of Female Menders in Woolen and Worsted Establishments.

What are some of the things which these summary figures show?

1. The arithmetic mean exceeds both the median and the mode [1] in each year. Skewness is, therefore, positive.

2. Both the average and the standard deviations, as well as the coefficients of dispersion based on them, tend to increase from year to year. That is, the average differences in rates when measured from the arithmetic mean tend to be larger both absolutely and relatively.

3. The lower quartile position in 1907 is essentially as high as the median in 1909. The range of difference in rates between the median and the upper quartile is more than double in 1910 what it is in 1907.

4. In both 1909 and 1910 there is a much more pronounced skew between the medians and the upper quartiles than in 1907 and 1908, the coefficients on the quartile measures being, respectively, + .21, + .13, + .40, + .31.

5. The wage-rates which the middle-half received varied as follows:

1907, from 12.07 to 16.32 or 4.25¢.
1908, from 11.48 to 17.77 or 6.29¢.
1909, from 10.14 to 16.61 or 6.47¢.
1910, from 11.05 to 18.52 or 7.47¢.

That is, the position of the lower quartile, with one exception, has fallen, and that of the upper quartile, with one exception, risen. While the average rate in 1910 is less than one half cent higher than in 1907, the wage of the person three fourths up in the scale is more than two cents higher.

6. The coefficient of dispersion based on the average deviation, and the coefficient of skewness based on the quartile measure are higher in 1909 and 1910 than in any

[1] A single mode is indeterminate in 1908 and 1910.

other of the years. Skewness indicates a healthy influence in wage conditions — a concentration above the arithmetic mean. On the other hand, the wide absolute and relative dispersions tend to counteract this.

Other detailed facts may be gleaned from a comparison of these summaries, but those given are sufficient to show their possibilities. It is generally not enough to speak in terms of averages when characterizing complex things. Deviations both as to amount and position are frequently vital and ought not to be ignored. By means of these an approach is scientific, since discrimination is made between things which by simple and undifferentiated criteria appear alike.

IV. Conclusion

In this chapter there have been outlined the meaning, measures and coefficients of dispersion and skewness and the methods by which they are computed. The mathematical side of the problem both in use of terms and in the tone of discussion has purposely been omitted or neglected with the thought that by so doing the topics would appeal to those who are without such training. It is hoped that this has not resulted in confusing those who habitually think in terms of mathematical symbols, or in sacrificing science to expediency. In the principles and the application that may be made of them the student and statistician are furnished with tools for the interpretation of everyday statistical facts. The discrimination which their use implies will serve as a safeguard against the serious error of failing to take account of differences and against the temptation to always speak in terms of averages — "an excuse for laziness."

References

Clark, Earle. — "The Horizontal Zero in Frequency Diagrams," in *Quarterly Publications of the American Statistical Association*, June, 1917, pp. 662–669.

Davenport, Eugene. — *Principles of Breeding*, Ch. XII, pp. 419–452.

Elderton, W. P. and E. M. — *Primer of Statistics*, Ch. IV, pp. 40–55.

King, W. I. — *Elements of Statistical Method*, Chs. XIII and XIV, pp. 141–167.

Mitchell; W. C. — "Methods of Presenting Statistics of Wages," *Publications of American Statistical Association*, Vol. IX, pp. 325–343.

Thorndike, E. L. — *Theory of Mental and Social Measurements*, Chs. III and IV, pp. 28–42, and 42–63, respectively.

Yule, G. U. — *An Introduction to the Theory of Statistics*, Ch. VIII, pp. 133–156, sections 1, 2, 5 to 11, 13 to 30.

Žižek, Franz. — *Statistical Averages*, Part III, Chs. I and II, pp. 251–255 and 256–270, respectively.

CHAPTER XII

COMPARISON — CORRELATION

I. Introduction

The preceding chapters, for the most part, have had to do with the preliminaries to comparison. These include units of measurements; coefficients of time, place, and condition; averages as types; measures and coefficients of dispersion and skewness; etc. The method of the discussion has been to consider first, loose, heterogeneous, and undifferentiated data, and, second, the means by which they are reduced and classified according to the logic of a clearly formulated statistical purpose. Everything has been directed toward quantitative comparison as the goal in statistical study.

II. The Meaning of Comparison and What It Implies Statistically

Comparison is made between things possessing common qualities. These may be of time, of place, or of condition. For instance, the accident rate in a given industry may be compared before and after the installation of safety devices. Comparison may extend to two industries operating at different places or under different conditions, the purpose being merely to record a quantitative difference. But comparisons are rarely made for this alone. Generally, a more or less definite purpose of establishing causal connection lies in the background. A specific inquiry is to determine

whether phenomena stand in the relation of cause and effect, or whether they are the result of a common cause. Whatever the purpose, a condition of comparison is that things compared have qualities in common.

The establishment of cause and effect relationships in economic studies offers great temptation and at the same time great difficulty. This is especially true when statistics are relied upon because so frequently they are incomplete, biased, and generally faulty. They are too often only seemingly exact. But whatever the tool, things compared are simply recorded experiences. These grow out of the facts of business, out of the observations of science, out of the records of history, etc., but are different for different people, for different times, and for different conditions. Their seeming unity and identity are, therefore, only relative and the order of cause and effect not implacable.

In actual life, business or otherwise, experiences grow out of environment variously interpreted. Variation at a given time and change over a period of time characterize the whole economic and business world. There are degrees of difference at the same time, between periods and over areas, but all traceable to a complex of causes. A given cause is not a homogeneous thing except when viewed in the broadest way. The effects which seem to follow from it do not come as an undifferentiated whole, but likewise as variations. Some come as coincidences, others as sequences spread over long or short periods. The assignment of cause and effect must be in keeping with the fact that a single cause is rarely found, and if found cannot be said always to give rise to a single effect. Both cause and effect [1] are in reality variates.

How true this is may be seen by briefly referring to some

[1] *Cf.* Hooker, R. H., "Correlation of the Marriage Rate with Trade," *Journal of the Royal Statistical Society*, Vol. 64, p. 485.

of the more common relationships among business phenomena. Stimulation of business is registered in bank clearings, but not all banks are equally affected. The effect upon the interest rate comes late and is far from being uniformly felt. Excessive issues of irredeemable paper currency ultimately result in a premium on gold and a general increase in prices, but not concurrently with the issue nor to the same degree in all countries or in different parts of the same country. It is only when a lack of confidence in the government develops that the premium is significant. The surplus reserves of banks are said currently to fix the call-loan interest rate. But not all loans, nor all banks nor customers, are affected at the same time and to the same degree. Somewhat later the effect of a marked surplus reserve is seen in the interest rate on 60- and 90-day commercial paper and on stock exchange collateral, but even then, not the same for every circumstance. Wholesale and retail prices fluctuate together, but the former fall first and rise first, retail prices following some distance behind. In this case cause and effect show themselves as a sequence. But neither all wholesale nor all retail prices respond in precisely the same way, nor is the response uniform from place to place nor from time to time. The effect of cotton prices on acreage is shown only from one cropping to another, and then not uniformly over the cotton area. Wages undoubtedly tend to rise with rising prices, but not coincidently, nor to the same degree in all trades. Other forms of labor remuneration, not included under the term "wages," as well as wages as generally understood, may actually fall during such periods when measured in terms of purchasing power. Business prosperity undoubtedly stimulates immigration but the cycle through which it passes begins later and is longer than is that for business prosperity, as indicated

by wholesale prices.[1] The relation is clearly that of a sequence. Moreover, the response is not an undifferentiated thing. No doubt, those most affected by pecuniary motives respond first and those later or not at all who are actuated differently. Again, general business prosperity exerts a greater influence than that which is limited and particular, but so-called *general* prosperity is far from uniform either for areas, for industries, or for people, etc.

Comparison, therefore, involves the pairing of things or events which are not identical in all particulars as to time, place, and condition. Causation in fact becomes contingency or correlation. A study of cause and effect, whether of coincidence or sequence, becomes largely a study of association. The idea that a given effect is the result of a specific cause and that there can be no other, or that the result must in the nature of the case be uniform and absolute, does not apply to business and economic phenomena. Causes never operate under exactly the same circumstances. Oneness of effect is only apparent, variation being evident the moment that the scale of measurement is reduced. When making comparison in economics or business, there is a tendency to attempt to safeguard oneself against error and criticism by introducing the proviso — *other things being equal.* But the "other things" are rarely if ever equal in actual life. To expect that an absolute cause will always result in an absolute effect or that the "other things" will automatically take care of themselves is futile. A realization of this fact will go a long way toward dispelling the tendency among business men and students to look for short cuts to success, as a result to the adoption of a rule-of-thumb formula, and

[1] Professor Persons finds immigration to be a business barometer when correlated with *yearly* wholesale prices. Had inter-annual correlations been worked out, the agreement, it is believed, would not have been simultaneous.

to expect that certain results will always follow an application of the *appropriate* rule of action.

Business does not go on indefinitely repeating itself in one unending round of sameness, and this fact is slowly being realized. The belief in the adequacy of the goad as a means of increasing output is slowly being dispelled. Those responsible for successful business are coming to believe that employees must be made to feel that they are a part of an organization, for only in this way is it possible to cut down the costs due to labor difficulties, rapid turnover of labor force, etc. In merchandising, competition is teaching that merely to place a commodity on the market and to sit back and wait for custom no longer suffice. Advertising in accordance with psychological principles is proving its power to bring out some unexpressed want or some new desire in its successes. But response to a campaign of advertising, for instance, is not unitary and absolute; it is diversified and varied. It is not unconditional and complete, but halting and partial. Variation characterizes this as it does all phenomena which involve the human element, whether viewed as cause or as effect. The tendency to look upon business and economic phenomena in a mechanistic manner, to expect a complete and narrow fulfilment of *the law* of cause and effect, must be dispelled. Just as soon as it is, the way is open for the operation of scientific method, not alone in the so-called scientific world, but in business at large. This is the method of discrimination, of the study of small differences, of acting in the light of facts properly interpreted, and of reducing them as classified knowledge into rules of guidance.

The rules to which facts point may be nothing more, for instance, than that it is unwise to market corn with high moisture content, since weight varies inversely with mois-

ture,[1] or to leave corn in leaky cars exposed to hot weather because both are conducive to the development of acidity, and acidity retards germination;[2] that a "bacon" hog can be produced; that corn grown from seed from ears 10 inches long has, on the average, longer ears than corn grown from seed of ears that are eight inches long;[3] that the prices of bonds with fixed interest rates vary inversely with general commodity price changes;[4] that a farm of less than forty acres in a certain district is economically undesirable;[5] that the milk production of cows increases up to at least six years of age and then falls off;[6] that there is a direct relation between fatigue and industrial accidents;[7] that accident rates tend to increase with expanding and to contract with falling business;[8] that twin offspring from twin parents in sheep production is more common than from parentage conforming to any other condition;[9] etc. Whatever they are and to whatever type

[1] *Bulletin* of the United States Department of Agriculture, No. 472, October, 1916, "Improved Apparatus for Determining the Test Weight of Grain, with a Standard Method of Making the Test." See curve on p. 4.

[2] *Bulletin* of the United States Department of Agriculture, No. 102, July, 1914, on "Acidity as a Factor in Determining the Degree of Soundness of Corn," pp. 12, 14, *passim.*

[3] "Type and Variability in Corn," *Bulletin* No. 119, University of Illinois Agricultural Experiment Station, October, 1907.

[4] Mitchell, Wesley C., *Business Cycles*, pp. 201–219, especially charts 23 and 24, pp. 206 and 207, respectively.

[5] *Bulletin* of the United States Department of Agriculture, No. 341, January, 1916, on "Farm Management Practice of Chester County, Pa.," pp. 56 ff.

[6] Holdaway, C. W., "Statistical Weighting for Age of Advanced Registry Cows," *The American Naturalist*, Vol. 50, No. 559, p. 681.

[7] "The Case for the Shorter Day," *Franklin O. Bunting* vs. *The State of Oregon*, Brief for the Defendant in Error, by Felix Frankfurter, Vol. 1, pp. 165–193.

[8] Mowbray, A. H., and Black, S. B., "Relation of Accident Frequency to Business Activity," in *Proceedings of the Casualty, Actuarial and Statistical Society of America*, Vol. 11, Pt. III, No. 6, May, 1916, pp. 418–426.

[9] Rietz, H. L., and Roberts, Elmer, "Degree of Resemblance of Parents and Offspring with Respect to Birth of Twins for Registered Shropshire Sheep," in *Journal of Agricultural Research*, Vol. IV, No. 6, September, 1915.

of business they apply, if they are arrived at as a result of a dispassionate study of facts in an attempt to determine association and correlation and not to *prove* the infallibility of some narrow cause-and-effect relationship, a clear advance is made in the use of statistical methods.

This has long been recognized. But what are facts, particularly *statistical* facts? Where are they, and what are the methods by which they may be used inside and outside of business? These are the questions which are now being asked and which it is the purpose of much that is written here to explain. Just as soon as business men and others dealing with economic science come to realize that rules of business cannot be read in the movements of heavenly bodies and traced out in some natural order, or divined by some occult formula, just so soon will real progress be made. It is not so much a question of *reading* the answer to a business problem as it is of understanding and applying facts to business. In no definite sense may the solution of business problems be found in a rule-of-thumb formula.

III. The Meaning of Correlation

If the establishment of causation in a narrow sense is impossible in economic and business science, since causes operate as variations and effects show themselves in the same way, it is unnecessary to conclude that cause and effect relationships in a larger sense cannot be measured. The problems are different and should be kept distinct. The first is the impossible task of establishing an absolute cause and an absolute effect; the latter is the problem of measuring correlation. Pearson makes the distinction clear in the following passage:

"When we vary the cause, the phenomenon changes, but not always to the same extent; it changes, but has variation in its

change. The less the variation in that change, the more nearly the cause defines the phenomena, the more closely we assert the association or the correlation to be. It is this conception of correlation between two occurrences embracing all relationships from absolute independence to complete dependence, which is the wider category by which we have to replace the old idea of causation. Everything in the universe occurs but once, there is no complete sameness of repetition. Individual phenomena can only be classified, and our problem turns on how far a group or class of like, but not absolutely same, things which we term 'causes' will be accompanied or followed by another group or class of like, but not absolutely same things which we term 'effects.'"[1]

What correlation, as thus distinguished from causation, means, is indicated in the quotations immediately following.

"When two quantities are so related that the fluctuations in one are in sympathy with fluctuations in the other, so that an increase or decrease of one is found in connection with an increase or decrease (or inversely) of the other, and the greater the magnitude of the changes in the one, the greater the magnitude of the changes in the other, the quantities are said to be *correlated.*"[2]

"The whole subject of correlation refers to that interrelation between separate characters by which they tend, in some degree at least, to move together. This relation is expressed in the form of a ratio. Thus, if an increase of one character is always followed by a corresponding and proportional increase in a related character, the correlation is said to be perfect and the ratio is 1. On the other hand, if an increase in one character is followed by a corresponding and proportional *decrease* in a related character, the correlation is said to be negative and the ratio is -1, or perfect negative correlation. Still again, if the characters in question are absolutely indifferent the one to the other, the correlation is said to be zero, indicating mere association under the law of independent probability, without causative relation of any kind."[3]

[1] Pearson, Karl, *The Grammar of Science*, p. 157.
[2] Bowley, A. L., *Elements of Statistics*, p. 316.
[3] Davenport, Eugene, *Principles of Breeding*, p. 453.

An experiment conducted by Professor Weldon [1] and carried further by Darbishire brings out clearly the meaning of correlation. Darbishire [2] found that by taking 12 dice and throwing them 1000 times, and counting the number that had four or more spots uppermost at each trial, he got the following distribution:

TABLE SHOWING THE DISTRIBUTION OF DICE WITH FOUR OR MORE SPOTS UPPERMOST IN 1000 THROWS (Darbishire)

RESULT OF THROW	FREQUENCY	RESULT OF THROW	FREQUENCY
0	0	7	179
1	3	8	129
2	15	9	64
3	55	10	11
4	110	11	2
5	208	12	1
6	223		

Another set of 1000 trials undoubtedly would have given a similar, but not necessarily the same, distribution.[3] Successive throws, after each of which all dice are returned to the receptacle and thrown again, are entirely distinct. There is no connecting link between them which makes them stand in the relation of cause and effect. This is shown in the following double-frequency or correlation table — Table A — where throws in pairs are tabulated.

[1] Weldon, W. F. R., "Inheritance in Animals and Plants," pp. 81–100, in *Lectures on the Method of Science*, edited by T. B. Strong, Oxford, 1906.

[2] Darbishire, A. D., "Some Tables for Illustrating Statistical Correlation," in *Memoirs and Proceedings of the Manchester Literary & Philosophical Society*, Vol. 51, No. 16, 1907.

[3] Cf. Weldon, W. F. R., *op. cit.*, for the results of three trials.

TABLE A

TABLE GIVING THE RESULTS OF 500 PAIRS OF THROWS OF 12 DICE WHEN ALL THOSE THROWN THE FIRST TIME WERE THROWN THE SECOND TIME [1]

			SECOND THROWS												
			0	1	2	3	4	5	6	7	8	9	10	11	12
		Total	→	1	9	24	57	112	101	94	62	31	6	2	1
	0	↓													
	1	2	—	—	—	—	—	1	—	—	1	—	—	—	—
	2	6	—	—	—	1	—	4	—	—	—	1	—	—	—
	3	31	—	—	—	1	4	7	8	5	4	1	1	—	—
	4	52	—	—	4	4	7	9	6	12	5	5	—	—	—
First	5	95	—	—	3	5	13	26	14	14	12	6	1	1	—
Throws	6	123	—	—	1	6	15	25	24	28	15	6	2	1	—
	7	87	—	—	1	5	7	16	22	15	13	6	1	—	1
	8	66	—	—	—	1	7	15	19	12	6	6	—	—	—
	9	33	—	1	—	1	2	9	7	6	6	—	1	—	—
	10	5	—	—	—	—	2	—	1	2	—	—	—	—	—
	11		—	—	—	—	—	—	—	—	—	—	—	—	—
	12		—	—	—	—	—	—	—	—	—	—	—	—	—

The data were secured as follows: Twelve dice were thrown a first time and the number having four or more spots uppermost counted. They were then all picked up, reshaken, and thrown again, those having four or more spots uppermost being again counted. This constituted a second time and completed the first pair of trials. Five hundred such pairs of trials, or one thousand separate throws in all, were then tabulated so that the figures on the vertical scale represented

[1] The order of the units in the ordinate scale is reversed in this instance from that usually followed.

the first count in each pair of trials and the figures on the horizontal the second count in each. For instance, the 14 in the 6th (vertical) and in the 5th row (horizontal) means that of the 500 pairs of trials, there were 14 in which the first throw of the trial gave 5 dice with 4 or more spots upward and in the second throw of the trial 6 dice with 4 or more spots uppermost. The figures in all other squares are similarly accounted for. The vertical totals give the distributions for the first throws; and the horizontal totals, the distributions for the second throws. The most probable number of dice showing 4 or more spots uppermost in a throw of twelve is six, but the number may be anything between zero and 12. The concentration at or near six in both totals shows this to be true for the 1000 separate throws.

Data in this form show no causal connection between the first and second throws in each pair. For instance, when there were seven dice in the first throws with 4 or more spots uppermost, there were from 2 to 12 with 4 or more in the second trials. Dispersion is equally noticeable in the opposite direction. When 8 fulfilled the conditions in the second trials, the corresponding numbers in the first throws varied from 1 to 9.

In order to connect or relate the two throws of each pair, Darbishire repeated the experiment, first leaving down and counting in the second throw of each pair one, then two, then three, etc., of the dice which previously had been stained red so as to distinguish them from the others. The experiment was continued until all of the 12 dice thrown in the first, were left down for the second throws. The results when 3, 5, and 10 dice were left down are given in Tables B, C, D, respectively. Correlation is shown graphically in Plate 26.

TABLE B

TABLE GIVING THE RESULTS OF 500 CONNECTED THROWS OF 12 DICE, IN EACH SECOND THROW OF WHICH 3 DICE WERE LEFT DOWN AND COUNTED [1]

			Second Throws												
			0	1	2	3	4	5	6	7	8	9	10	11	12
		Total	→	2	7	31	55	82	111	108	71	25	7	1	—
	0	↓													
	1	2	—	—	—	—	1	—	1	—	—	—	—	—	—
	2	7	—	—	—	—	—	—	6	1	—	—	—	—	—
	3	20	—	1	1	5	2	2	4	5	—	—	—	—	—
	4	64	—	—	1	8	6	21	16	6	6	—	—	—	—
First	5	92	—	—	4	3	12	15	23	22	9	3	1	—	—
Throws	6	123	—	1	—	10	16	17	23	28	22	5	1	—	—
	7	97	—	—	1	4	9	17	18	24	16	5	3	—	—
	8	54	—	—	—	1	5	6	10	14	8	7	2	1	—
	9	30	—	—	—	—	4	3	9	6	6	2	—	—	—
	10	10	—	—	—	—	—	1	1	1	4	3	—	—	—
	11	1	—	—	—	—	—	—	—	1	—	—	—	—	—
	12		—	—	—	—	—	—	—	—	—	—	—	—	—

In each pair of trial throws, in which one or more of the dice is left on the board and counted in the second throw, there is a common element. That is, the first is in part a cause of the second, exerting an influence in proportion to its size. But the distributions in none of the cases, if the trials were repeated, would necessarily follow the order here given. When the two throws of the pairs are independent, there is little or no correlation present; when the second throw is simply the first counted as the second, correlation

[1] See note to Table A.

TABLE C

Table Giving the Results of 500 Connected Throws of 12 Dice, in Each Second Throw of Which 5 Dice Were Left Down and Counted[1]

			Second Throws												
			0	1	2	3	4	5	6	7	8	9	10	11	12
		Total	→		11	20	54	93	112	118	60	21	9	2	—
	0	↓													
	1	2	—	—	—	—	1	1	—	—	—	—	—	—	—
	2	11	—	—	3	1	5	1	1	—	—	—	—	—	—
	3	26	—	—	3	3	8	4	4	4	—	—	—	—	—
	4	69	—	—	3	6	9	21	14	10	5	1	—	—	—
First Throws	5	83	—	—	—	4	11	23	21	15	9	—	—	—	—
	6	109	—	—	1	3	9	18	27	29	16	3	2	1	—
	7	95	—	—	1	2	5	14	24	28	10	7	4	—	—
	8	63	—	—	—	1	5	9	10	18	14	4	2	—	—
	9	31	—	—	—	—	—	2	9	13	4	3	—	—	—
	10	10	—	—	—	—	1	—	2	—	2	3	1	1	—
	11	1	—	—	—	—	—	—	—	1	—	—	—	—	—
	12		—	—	—	—	—	—	—	—	—	—	—	—	—

is perfect and positive. The other trials show correlation between zero and + 1.

But the presence of a high degree of correlation cannot logically be said to *prove* the relation between two phenomena.[2] Causes never operate at different times under exactly the same conditions, and the effects that flow from them are not always and necessarily the same. Duplication

[1] See note to Table A.

[2] Cf. Hooker, "Correlation of the Marriage Rate with Trade," *Journal of the Royal Statistical Society*, Vol. 64, p. 485.

TABLE D

Table Giving the Results of 500 Connected Throws of 12 Dice, in the Second Throws of Which 10 Dice Were Left Down and Counted [1]

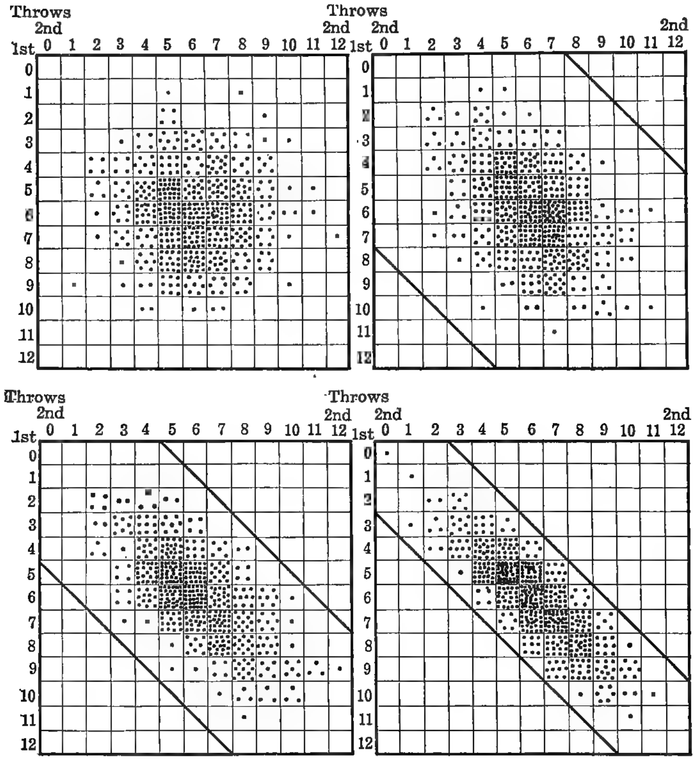
			Second Throws												
		500	0	1	2	3	4	5	6	7	8	9	10	11	12
		Total →	1	2	7	24	55	93	111	100	64	31	11	1	
		↓													
	0	1	1	—	—	—	—	—	—	—	—	—	—	—	—
	1	1	—	1	—	—	—	—	—	—	—	—	—	—	—
	2	7	—	—	2	5	—	—	—	—	—	—	—	—	—
	3	24	—	1	3	8	9	3	—	—	—	—	—	—	—
First	4	55	—	—	2	10	18	19	6	—	—	—	—	—	—
Throws	5	110	—	—	—	1	24	43	32	10	—	—	—	—	—
	6	93	—	—	—	—	4	22	37	24	6	—	—	—	—
	7	96	—	—	—	—	—	6	27	39	19	5	—	—	—
	8	60	—	—	—	—	—	—	9	17	24	9	1	—	—
	9	42	—	—	—	—	—	—	—	10	14	11	7	—	—
	10	10	—	—	—	—	—	—	—	—	1	6	2	1	—
	11	1	—	—	—	—	—	—	—	—	—	—	1	—	—
	12		—	—	—	—	—	—	—	—	—	—	—	—	—

of the conditions under which causes operate will not necessarily duplicate the effects. "Duplication" after all in any way except as approximation is impossible in actual life. A measure of correlation is a statement of probabilities, the reliability of which is determined by the degree to which the samples represent the whole "population," and the conditions under which the samples are taken the range of conditions. "It does not prove anything. It merely suggests an hypothesis as regards causation within a particular sphere.

[1] See note to Table A.

Second Throws Independent of First Throws.

Second Throws Dependent on First Throws — 5 Dice in Common.



Second Throws Dependent on First Throws — 8 Dice in Common.

Second Throws Dependent on First Throws — 10 Dice in Common.

PLATE 26

Graphic Figures Illustrating Correlation by Means of 500 Pairs of Throws of Dice.

The investigator must go to the facts themselves for his scientific hypothesis."[1]

How nearly economic and business phenomena remain homogeneous for any appreciable period, even in an approximate sense, is always problematical. The forces affecting them are always in a state of flux governed as they are by population composition, state of trade, distribution of wealth, custom, fad, fashion, prejudice, etc. The whole range of human reaction is exhibited in more or less degree. Statistics under such circumstances often reveal a partial story, are not comparable from time to time and from place to place, and taken alone constitute a weak and uncertain base upon which to build a cause-and-effect structure. Statistical studies in correlation should be made in the light of these facts. Again, statistics and statistical methods, used as tools in induction, are serviceable only to the degree to which they are properly employed.

1. *Preliminaries to Correlation Studies (Historical Series)*

When historical series are to be compared it is often serviceable, as a preliminary step, to plot them side by side in order to determine whether increases or decreases in one tend to conform to increases or decreases (or the inverse) in the other. Frequently, this in itself is sufficient to suggest correlation.[2] But the graphic method, though suggestive, is neither proof nor measure of correlation.[3] It does not give a quantitative measure of the degree of resemblance, and this is what is sought.

[1] Brown, William, *The Essentials of Mental Measurements*, p. 132; see also, Sigwart, C., *Logic*, Vol. II, p. 502.

[2] Bowley, *Measurement of Groups and Series*, p. 84.

[3] Cf. Persons, Warren M., "The Correlation of Economic Statistics," *Proceedings of the American Statistical Association*, Vol. XII, New Series, December, 1910, pp. 287–322.

Experiences and business facts are associated either as coincidences or as sequences. What is sought is some means of foretelling the consequences of a given line of action, of discounting the future.[1] Sometimes the full effects of a cause are not felt for a period of time, as for instance, price changes on wages, bank clearings on wholesale prices, unemployment on sickness among trade unionists,[2] etc. Moreover, the period of delay is not uniform. In some instances, sufficient time for a set of causes to exert its full effect requires years,[3] in other cases, days. An approximation to the correct time which one historical series lags behind another may be made by a series of graphic tests. This, however, is simply following the trial and error method. It is possible, by the use of quantitative measures, to discover the period in which there is most complete correlation and to plot data in this form. What these measures are and what they mean are the subjects of the next section.

Moreover, in correlating historical series, it is frequently necessary to distinguish between short- and long-time changes. Two phenomena may be correlated when long or secular changes are considered, but be entirely disassociated for short or cyclic changes. Or, for short periods two series may move together, but show no connection for an extended period. The question is to decide which are to be correlated.

In an earlier chapter[4] the method of smoothing historical series by means of moving averages was described. In

[1] Persons, Warren M., "Construction of a Business Barometer Based upon Annual Data," *The American Economic Review*, December, 1916, p. 755.

[2] Ashton, T. S., *Economic Journal*, September, 1916, p. 396.

[3] Moore, H. L., *Economic Cycles: their Law and Cause*, Ch. V, *passim*.

[4] *Supra*, pp. 229–230.

Table F and in Plate 27, giving bank note circulation [1] of chartered Canadian banks and receipts of wheat at Fort William and Port Arthur, Canada, this device is more fully illustrated and its relation to methods of determining correlation shown.

[1] "The redemption system, besides making currency inflation impossible, results also in what is commonly called 'elasticity,' by which is meant capacity to expand and contract in automatic response to the country's need of currency. Canada, like every other country, at certain seasons of the year makes use of more currency, or hand-to-hand money, than at other seasons. This currency is supplied by the banks. If they were not permitted to furnish it in the form of their own notes, they would be obliged to furnish it in the form of lawful or legal tender money, and would at the same time be compelled to restrict their loans in order that they might reduce their liabilities, the loss of the legal tender money having by so much reduced their cash reserve. Since the Canadian banks, however, meet the seasonal needs for currency by the issue of notes, their liabilities are not changed, for their deposits decline by as much as their notes increase. It is clear that if a depositor draws $1000 from his bank and receives $1000 in the notes of the bank, the liabilities of the bank have not been affected. It has simply converted a deposit liability into a note liability. Its reserve of legal tender money having been untouched, it is under no necessity to reduce its loans. It follows that since a Canadian bank is able to supply its depositors' needs for cash with its own bank notes, it can do so without being compelled to lessen its usefulness to the community as a lender of money." Johnson, Joseph French, *The Canadian Banking System*, pp. 61–62.

TABLE E

TABLE SHOWING THE LONG-TIME OR SECULAR CHANGES IN NOTE CIRCULATION OF CHARTERED CANADIAN BANKS AND WHEAT RECEIPTS AT FORT WILLIAM AND PORT ARTHUR, CANADA, 1909–1913

YEAR AND MONTH	NOTES IN CIRCULATION (In millions of dollars)	DEVIATIONS FROM AVERAGE CIRCULATION x	DEVIATIONS SQUARED x^2	WHEAT RECEIVED (In millions of bushels)	DEVIATIONS FROM AVERAGE RECEIPT y	DEVIATIONS SQUARED y^2	PRODUCT OF DEVIATIONS (x) and (y)
a	*b*	*c*	*d*	*e*	*f*	*g*	*h*
Total	95 av.		13406	8.0 av.		3805.2	+ 4799.4
1909							
Jan.	73	– 22	484	2.1	– 5.9	34.8	+ 129.8
Feb.	68	– 27	729	1.6	– 6.4	41.0	+ 172.8
Mar.	71	– 24	576	3.4	– 4.6	21.2	+ 110.4
Apr.	73	– 22	484	3.9	– 4.1	16.8	+ 90.2
May	71	– 24	576	1.6	– 6.4	41.0	+ 153.6
June	72	– 23	529	.6	– 7.4	54.8	+ 170.2
July	74	– 21	441	1.5	– 6.5	42.3	+ 136.5
Aug.	74	– 21	441	.2	– 7.8	60.8	+ 163.8
Sept.	82	– 13	169	11.1	+ 3.1	9.6	– 40.3
Oct.	91	– 4	16	17.0	+ 9.0	81.0	– 36.0
Nov.	92	– 3	9	15.1	+ 7.1	50.4	– 21.3
Dec.	90	– 5	25	6.8	– 1.2	1.4	+ 6.0
1910							
Jan.	81	– 14	196	2.7	– 5.3	28.1	+ 74.2
Feb.	76	– 19	361	1.7	– 6.3	39.7	+ 119.7
Mar.	81	– 14	196	2.8	– 5.2	27.0	+ 72.8
Apr.	82	– 13	169	4.2	– 3.8	14.4	+ 49.4
May	81	– 14	196	4.5	– 3.5	12.3	+ 49.0
June	82	– 13	169	2.2	– 5.8	33.6	+ 75.4

TABLE E *Continued*

Year and Month	Notes in Circulation (In millions of dollars)	Deviations from Average Circulation x	Deviations Squared x^2	Wheat Received (In millions of bushels)	Deviations from Average Receipt y	Deviations Squared y^2	Product of Deviations (x) and (y)
a	*b*	*c*	*d*	*e*	*f*	*g*	*h*
1910							
July	84	− 11	121	2.8	− 5.2	27.0	+ 57.2
Aug.	85	− 10	100	1.5	− 6.5	42.3	+ 6.5
Sept.	90	− 5	25	8.5	+ .5	.3	− 2.5
Oct.	97	+ 2	4	18.6	+ 10.6	112.4	+ 21.2
Nov.	99	+ 4	16	13.3	+ 5.3	28.1	+ 21.2
Dec.	95	0	0	6.1	− 1.9	3.6	.0
1911							
Jan.	86	− 9	81	1.0	− 7.0	49.0	+ 63.0
Feb.	82	− 13	169	1.0	− 7.0	49.0	+ 91.0
Mar.	86	− 9	81	4.2	− 3.8	14.4	+ 34.2
Apr.	90	− 5	25	5.2	− 2.8	7.8	+ 14.0
May	87	− 8	64	3.5	− 4.5	20.3	+ 36.0
June	90	− 5	25	3.5	− 4.5	20.3	+ 22.5
July	93	− 2	4	4.5	− 3.5	12.3	+ 7.0
Aug.	94	− 1	1	1.7	− 6.3	39.7	+ 6.3
Sept.	100	+ 5	25	5.7	− 2.3	5.3	− 11.5
Oct.	107	+ 12	144	19.3	+ 11.3	127.7	+ 135.6
Nov.	111	+ 16	256	19.9	+ 11.9	141.6	+ 190.4
Dec.	110	+ 15	225	16.4	+ 8.4	70.6	+ 126.0
1912							
Jan.	101	+ 6	36	6.9	− 1.1	1.2	− 6.6
Feb.	93	− 2	4	6.7	− 1.3	1.7	+ 2.6
Mar.	98	+ 3	9	5.8	− 2.2	4.8	− 6.6
Apr.	102	+ 7	49	2.7	− 5.3	28.1	− 37.1
May	101	+ 6	36	9.7	+ 1.7	2.9	+ 10.2
June	103	+ 8	64	6.6	− 1.4	2.0	− 11.2

TABLE E *Continued*

Year and Month	Notes in Circulation (In millions of dollars)	Deviations from Average Circulation x	Deviations Squared x^2	Wheat Received (In millions of bushels)	Deviations from Average Receipt y	Deviations Squared y^2	Product of Deviations (x) and (y)
a	*b*	*c*	*d*	*e*	*f*	*g*	*h*
1912							
July	105	+ 10	100	5.4	− 6.2	6.8	− 26.0
Aug.	104	+ 9	81	3.1	− 4.6	24.0	− 44.1
Sept.	107	+ 12	144	2.7	− 5.3	28.1	− 63.6
Oct.	114	+ 19	361	19.6	+ 11.6	134.6	+ 220.4
Nov.	120	+ 25	625	27.6	+ 19.6	384.2	+ 490.0
Dec.	120	+ 25	625	15.0	+ 7.0	49.0	+ 175.0
1913							
Jan.	110	+ 15	225	12.1	+ 4.1	16.8	+ 61.5
Feb.	101	+ 6	38	4.1	− 3.9	15.2	− 23.4
Mar.	108	+ 13	169	2.4	− 5.6	31.4	− 72.8
Apr.	106	+ 11	121	2.7	− 5.3	28.1	− 58.3
May	105	+ 10	100	10.2	+ 2.2	4.8	+ 22.0
June	108	+ 13	169	5.5	− 2.5	6.3	− 32.5
July	108	+ 13	169	4.3	− 3.7	13.7	− 48.1
Aug.	109	+ 14	196	1.3	− 6.7	44.9	− 93.8
Sept.	114	+ 19	361	18.1	+ 10.1	102.0	+ 191.9
Oct.	124	+ 29	841	37.5	+ 29.5	870.3	+ 855.5
Nov.	127	+ 32	1024	30.9	+ 22.9	524.4	+ 732.8
Dec.	122	+ 27	729	17.9	+ 9.9	98.0	+ 267.3

TABLE F

TABLE SHOWING THE SHORT-TIME OR CYCLIC CHANGES IN NOTE CIRCULATION OF CANADIAN CHARTERED BANKS AND WHEAT RECEIPTS AT FORT WILLIAM AND PORT ARTHUR, CANADA, 1909–1913

YEAR AND MONTH	NOTES IN CIRCULATION (In millions of dollars)	MOVING AVERAGE (13 MONTH CYCLE)	DEVIATIONS FROM MOVING AVERAGE x	DEVIATIONS SQUARED x^2	WHEAT RECEIPTS (In millions of bushels)	MOVING AVERAGE (13 MONTH CYCLE)	DEVIATIONS FROM MOVING AVERAGE y	DEVIATIONS SQUARED y^2	PRODUCT OF (x) and (y)
a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*
Total	95 av.			1912.0	8.0 av.			1994.2	+ 1541.9
1909									
Jan.	73	—	—	—	2.1	—	—	—	—
Feb.	68	—	—	—	1.6	—	—	—	—
Mar.	71	—	—	—	3.4	—	—	—	—
Apr.	73	—	—	—	3.9	—	—	—	—
May	71	—	—	—	1.6	—	—	—	—
June	72	—	—	—	.6	—	—	—	—
July	74	77.8	− 3.8	14.4	1.5	5.3	− 3.8	14.4	+ 14.4
Aug.	74	78.1	− 4.1	16.8	.2	5.2	− 5.0	25.0	+ 20.5
Sept.	82	79.1	+ 2.9	8.4	11.1	5.3	+ 5.8	33.6	+ 16.8
Oct.	91	79.9	+ 11.1	123.2	17.0	5.3	+ 11.7	136.9	+ 129.9
Nov.	92	80.5	+ 11.5	132.2	15.1	5.4	+ 9.7	94.1	+ 111.6
Dec.	90	81.4	+ 8.6	74.0	6.8	5.4	+ 1.4	1.9	+ 12.0
1910									
Jan.	81	82.3	− 1.3	1.7	2.7	5.6	− 2.9	8.4	+ 3.8
Feb.	76	83.1	− 7.1	50.4	1.7	5.6	− 3.9	15.2	+ 27.7
Mar.	81	84.4	− 3.4	11.6	2.8	6.2	− 3.4	11.6	+ 11.6
Apr.	82	85.5	− 3.5	12.3	4.2	6.8	− 2.6	6.8	+ 9.1
May	81	86.1	− 5.1	26.0	4.5	6.5	− 2.0	4.0	+ 10.2
June	82	86.4	− 4.4	19.4	2.2	5.8	− 3.6	13.0	+ 15.8

TABLE F *Continued*

Year and Month	Notes in Circulation (In millions of dollars)	Moving Average (13 Month Cycle)	Deviations from Moving Average x	Deviations Squared x^2	Wheat Receipts (In millions of bushels)	Moving Average (13 Month Cycle)	Deviations from Moving Average y	Deviations Squared y^2	Product of (x) and (y)
a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*
1910									
July	84	86.1	− 2.1	4.4	2.8	5.4	− 2.6	6.8	+ 5.5
Aug.	85	86.2	− 1.2	1.4	1.5	5.3	− 3.8	14.4	+ 4.6
Sept.	90	87.0	+ 3.0	9.0	8.5	5.5	+ 3.0	9.0	+ 9.0
Oct.	97	87.7	+ 9.3	86.5	18.6	5.6	+ 13.0	169.0	+ 12.1
Nov.	99	88.1	+ 10.9	118.8	13.3	5.6	+ 7.7	59.3	+ 83.9
Dec.	95	88.8	+ 6.2	38.4	6.1	5.5	+ .6	.4	+ 3.7
1911									
Jan.	86	89.6	− 3.6	13.0	1.0	5.8	− 4.8	23.0	+ 17.3
Feb.	82	90.4	− 8.4	70.6	1.0	5.7	− 4.7	22.1	+ 39.5
Mar.	86	91.5	− 5.5	30.3	4.2	6.0	− 1.8	3.2	+ 9.9
Apr.	90	92.8	− 2.8	7.8	5.2	6.8	− 1.6	2.6	+ 4.5
May	87	93.9	− 5.9	34.8	3.5	6.9	− 3.4	11.6	+ 20.1
June	90	94.8	− 4.8	23.0	3.5	7.2	− 3.7	13.7	+ 17.8
July	93	95.2	− 2.2	4.8	4.5	7.2	− 2.7	7.3	+ 5.9
Aug.	94	95.8	− 1.8	3.2	1.7	7.7	− 6.0	36.0	+ 10.8
Sept.	100	97.0	+ 3.0	9.0	5.7	8.1	− 2.4	5.8	− 7.2
Oct.	107	98.2	+ 8.8	77.4	19.3	7.9	+ 11.4	130.0	+ 100.3
Nov.	111	99.1	+ 11.9	141.6	19.9	8.3	+ 11.6	134.6	+ 138.0
Dec.	110	100.3	+ 9.7	94.1	16.4	8.5	+ 7.9	62.4	+ 76.6
1912									
Jan.	101	101.5	− .5	.3	6.9	8.7	− 1.8	3.2	+ .9
Feb.	93	102.0	− 9.0	81.0	6.7	8.6	− 1.9	3.6	+ 17.1
Mar.	98	103.4	− 5.4	29.2	5.8	8.6	− 2.8	7.8	+ 15.1
Apr.	102	104.4	− 2.4	5.8	2.7	9.7	− 7.7	49.0	+ 16.8

TABLE F *Continued*

Year and Month	Notes in Circulation (In millions of dollars)	Moving Average (13 Month Cycle)	Deviations from Moving Average x	Deviations Squared x^2	Wheat Receipts (In millions of bushels)	Moving Average (13 Month Cycle)	Deviations from Moving Average y	Deviations Squared y^2	Product of (x) and (y)
a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*
1912									
May	101	105.4	− 4.4	19.4	9.7	10.3	− .6	.4	+ 2.6
June	103	106.0	− 3.0	9.0	6.6	10.0	− 3.4	11.6	+ 10.2
July	105	106.0	− 1.0	1.0	5.4	9.6	− 4.2	17.6	+ 4.2
Aug.	104	106.0	− 2.0	4.0	3.1	9.4	− 6.3	39.7	+ 12.6
Sept.	107	107.2	− .2	.1	2.7	9.1	− 6.4	41.0	+ 1.3
Oct.	114	107.8	+ 7.2	51.8	19.6	8.9	+ 11.7	136.9	+ 84.2
Nov.	120	108.0	+ 12.0	144.0	27.6	9.4	+ 18.2	331.2	+ 218.4
Dec.	120	108.5	+ 11.5	132.3	15.0	9.1	+ 5.9	34.8	+ 67.8
1913									
Jan.	110	108.9	+ 1.1	1.2	12.1	8.9	− 3.2	10.2	− 5.5
Feb.	101	109.2	− 8.2	67.2	4.1	8.6	− 4.5	20.3	+ 36.9
Mar.	108	110.0	− 2.0	4.0	2.4	9.8	− 7.4	54.8	+ 14.8
Apr.	106	111.3	− 5.3	28.1	2.7	12.5	− 9.8	96.0	+ 51.9
May	105	112.4	− 7.4	54.8	10.2	13.3	− 3.1	9.6	+ 22.9
June	108	112.5	− 4.5	20.3	5.5	12.6	− 7.1	50.4	+ 32.0
July	108				4.3				
Aug.	109				1.3				
Sept.	114				18.1				
Oct.	124				37.5				
Nov.	127				30.9				
Dec.	122				17.9				

For both note circulation and wheat receipts the short-time or cyclic changes seem to be approximately 13 months

in length.[1] These are removed by calculating moving averages on this wave length for both series, as given in columns *c* and *g*, respectively, in Table F. Graphically, they are shown by the smooth solid lines reflecting the trends in Plate 27. Their general direction in both cases is the same and over the whole period both phenomena show an unmistakable but somewhat different increase.

Diverting attention from the long- to the short-time movements of the two curves, regularity is observed in both. The movements tend to change together — *i.e.* increases and decreases in one roughly correspond to increases and decreases in the other. These cyclic changes — that is, the current differences from the trends or moving averages — are shown in columns *d* and *h* in Table F and graphically in Plate 28, where the differences are plotted as plus or minus deviations from a base or zero (no change) line. This illustration has the advantage of concentrating attention on the short-time fluctuations and of ignoring the long-time change.

In the example chosen, the relationship is one of coincidence.[2] The cause of increased circulation is in part the necessity of a circulating medium with which to move the crops. But crop harvesting and moving are largely the results of conditions of growth, ripening, lack of storage facilities at place of production, desire to sell at time of harvesting, etc. Seasonal influences are dominant and are reflected in bank circulation. These may be said to be a *cause* of increased but not of decreased circulation. The *cause* of the latter is the peculiarity of the banking system

[1] The use of a 13-months' cycle emphasizes the month repeated, but makes it possible to assign the moving mean to the middle item — the seventh. If 12 months — *i.e.* 12 values — had been used, the resulting mean would have fallen half-way between the sixth and seventh items.

[2] That is, within the period — a month in this case.

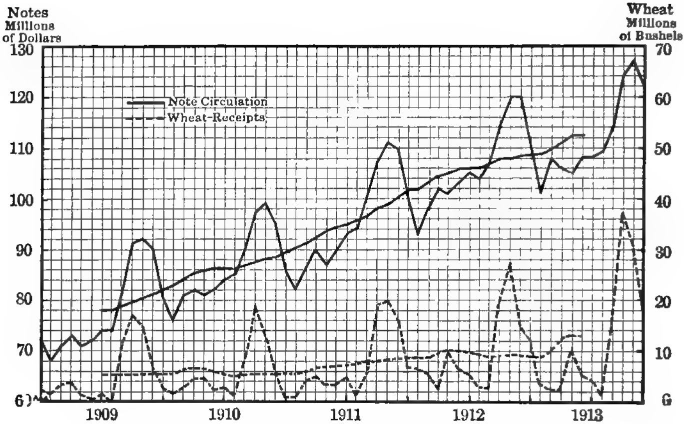


PLATE 27

Curves Showing Long-time or Secular Changes.

(Note Circulation of Canadian Chartered Banks, and Wheat Receipts at Fort William and Port Arthur, Canada, by Months, 1909–1913.)

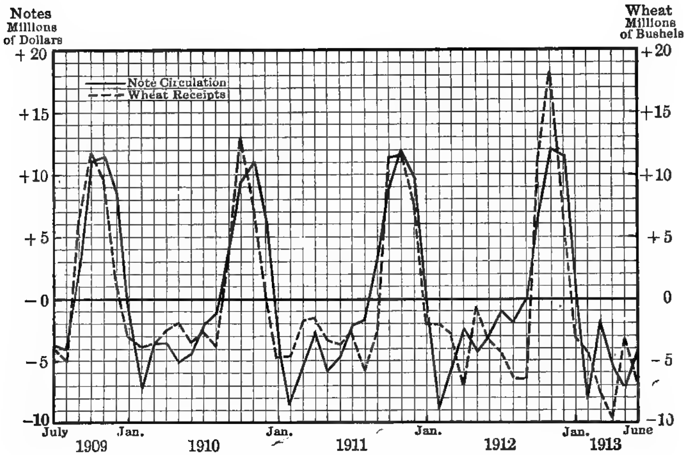


PLATE 28

Curves Showing Short-time or Cyclic Changes.

(Note Circulation of Canadian Chartered Banks, and Wheat Receipts at Fort William and Port Arthur, Canada, by Months, 1909–1913.)

which requires notes to be redeemed when no longer necessary. Demand for a circulating medium is due in part to wheat movement, but not solely so. Causation, in fact, becomes correlation. Both phenomena are related, but one is not the sole cause of the other.

How nearly these phenomena are related is suggested but not measured by the graphic method. The most common *measure* is the Pearsonian coefficient, developed by Sir Francis Galton and perfected by Karl Pearson in his studies of heredity. It has since become the tool of biometricians,[1] zoölogists,[2] breeders,[3] psychologists,[4] and economists.[5] Its latest development in the economic field is in the study of crises [6] and in the formation of a business barometer.[7] The remaining part of the chapter is devoted to explaining this measure and to showing its application to both historical and frequency series.

[1] See the journal *Biometrika* and the writings of Sir Francis Galton, Karl Pearson, C. B. Davenport, H. M. Vernon, *et al.*

[2] Among the leading is Harris, J. A., of the Carnegie Institution of Washington, D. C. See his "An Outline of Current Progress in the Theory of Correlation and Contingency," in *American Naturalist*, January, 1916, Vol. L, pp. 53–64.

[3] Davenport, Eugene, *The Principles of Breeding*, New York, 1907.

[4] Thorndike, E. L., *Mental and Social Measurements*, New York, 1913; Brown, William, *The Essentials of Mental Measurement*, Cambridge (England), 1911; Whipple, Guy M., *Manual of Mental and Physical Tests*, Baltimore, 1914.

[5] Hooker, R. H., *op. cit.;* Yule, *Introduction to Theory of Statistics*, London, 1911; Bowley, A. L., *Measurement of Groups and Series*, London, 1903; Elderton, W. Palin, *Frequency-curves and Correlation*, London, 1906 (?); Persons, W. M., "The Correlation of Economic Statistics," *Publications of the American Statistical Association*, Vol. XII, December, 1910, pp. 287-322.

[6] Moore, H. L., *Economic Cycles: Their Law and Cause*, New York, 1914.

[7] Persons, Warren M., "The Construction of a Business Barometer Based upon Annual Data," in *American Economic Review*, December, 1916, pp. 739–769.

2. *The Pearsonian Coefficient of Correlation*

Karl Pearson's coefficient of correlation is denoted by the formula, $r = \frac{\Sigma(xy)}{n\sigma_1\sigma_2}$[1], in which the x's are the series of deviations from the arithmetic mean of one series, and the y's the corresponding deviations from the arithmetic mean in the other series. The sign Σ stands for the algebraic sum of the products of the x's and y's. n refers to the number of pairs of items, and σ_1 and σ_2 to the respective standard deviations of the two series. The development of the formula gives values varying from -1 through 0 to $+1$.[2] If $r = +1$, correlation is perfect and positive — that is, *large values* in the first of two phenomena are associated with large values in the second. If $r = -1$, correlation is perfect and negative or inverse — that is, *large* (or *small*) values in the first of two phenomena are associated with *small* (or *large*) values in the second. If $r = 0$, no correlation exists, changes in the two phenomena being indifferent.[3]

The formula for the "coefficient was found by assuming that a large number of independent causes operate upon each of the two series x and y, producing normal distribution in both cases. Upon the assumption that the set of causes operating upon the series x is *not independent* of the set of causes operating upon the series y, the value $r = \frac{\Sigma xy}{n\sigma_1\sigma_2}$ is obtained. This value becomes zero when the operating causes are absolutely independent."[4]

[1] For the method by which this formula is derived, see Yule, G. Udny, *Introduction to the Theory of Statistics*, pp. 168–174.

[2] Proof in Bowley, A. L., *Elements of Statistics*, p. 319.

[3] Yule, *op. cit.*, p. 175.

[4] Persons, Warren M., "The Correlation of Economic Statistics," *Publications American Statistical Association*, December, 1910, pp. 298–299; Bowley, A. L., *Elements of Statistics*, pp. 316–317.

(1) Application of the Coefficient of Correlation to Historical Series

In the historical series — note circulation and wheat receipts — there are two movements that may be correlated. First, the long-time or secular changes;[1] and second, the short-time or cyclic changes. The latter, from the graphic representation, appear to move in unison and to stand in a causal relationship. The coefficient for the secular trend is calculated from the original rather than from the smoothed data, inasmuch as both long- and short-time changes are correlated positively. Had the secular trends been positively (or negatively) correlated and the periodic or cyclic changes negatively (or positively) correlated, it would have been necessary to use the moving averages. Even then difficulties would have arisen. As Bowley says:

> "If we take two things which are absolutely disconnected, except that they are both phenomena arising in the progress of society, and work out the coefficient by the straightforward rule, we shall find there is some correlation. If two curves have short fluctuations which are correlated, but opposite symptoms, then owing to the symptom apart from the fluctuations there would be negative correlation, while owing to the fluctuations apart from the symptom there would be positive correlation; and when both are taken into account the correlation may be positive, zero, or negative."[2]

But such is not the case in the example taken. In Table E, columns *c* and *f* give the monthly deviations — positive and negative — of note circulation and wheat receipts from their respective averages, 1909–1913. The respective

[1] Bowley calls them "symptomatic" changes. *Measurement of Groups and Series*, pp. 75–77.

[2] Bowley, A. L., *Measurement of Groups and Series*, p. 83.

standard deviations, computed according to the formula, $\sqrt{\frac{\Sigma d^2}{n}}$, are 14.9 and 7.96. The algebraic sum of the products of the deviations in the note series (x) and the wheat series (y) is found in column h, and equals $+\ 4799.4$. The coefficient of correlation r, by the formula, $\frac{\Sigma(xy)}{n\sigma_1\ \sigma_2}$, is $\frac{+\ 4799.4}{60 \times 14.9 \times 7.96}$ or $+\ 0.674$. That is, the correlation is positive and high. The probable error of the coefficient of correlation[1] based on the formula $.6745\ \frac{1 - r^2}{\sqrt{n}} = \pm\ .048$. The coefficient is 14 times the probable error and is therefore significant.[2]

The short-time or cyclic changes even on an inspection of the graphic figure show correlation. The degree of correlation may be measured by a slight modification of the Pearsonian coefficient used by Hooker,[3] Bowley,[4] Moore,[5] and others. It is employed in the series taken as an example. The deviations rather than being measured from the averages of the respective series are computed as given in Table F, columns d and h, by taking the current differences of the two series from their respective averages. These are squared as in columns e and i, as bases for computing the standard deviations. The products of the deviations in the

[1] See the discussion of the relation of the Probable Error to the normal curve of error distribution. The precise meaning of Probable Error of r is discussed by Bowley, *op. cit.*, pp. 88–90.

[2] Bowley says that r must be at least 6 times the Probable Error to be significant. Bowley, A. L., *Elements of Statistics*, p. 320. But significance can be attached to P. E. only on the assumption of a normal distribution.

[3] Hooker, R. H., "On the Correlation of the Marriage-rate with Trade," *Journal Royal Statistical Society*, Vol. LXIV, p. 486.

[4] Bowley, A. L., *Measurement of Groups and Series*, pp. 82–88.

[5] Moore, H. L., *Economic Cycles: Their Law and Cause*, Ch. V, *passim*.

x series and those in the y series are given in column j. Using the formula $r = \frac{\Sigma(xy)}{n\,\sigma_1\,\sigma_2}$, and inserting the values $\frac{+\,1541.9}{48 \times 6.45 \times 6.31}$, the coefficient of correlation is $+\,0.789$ with a probable error of $\pm\,.037$. That is, the correlation is positive and high and the probable error significant.[1]

In the example chosen it seems unnecessary to lag one series behind the other and to determine the correlation for various periods. Where the effect of a cause is not immediate, this is necessary. Recently, in two valuable studies related series have been lagged different periods and the coefficients calculated. Professor Moore, in correlating pig-iron production and yield of crops, says:

"If . . . we correlate them for lags of various intervals, we shall find it possible to determine the lag that will give the maximum coefficient of correlation, and this particular value of the lag we may then regard as the interval of time required for the cycles in the crops to produce their maximum effect upon the cycles of the activity of industry. When the calculation of the coefficients of correlation is made according to this plan, it is found that for a lag

Of zero years,	$r = .625$;
Of one year,	$r = .719$;
Of two years,	$r = .718$;
Of three years,	$r = .697$;
Of four years,	$r = .572$.

It is clear, therefore, that the cycles in the yield per acre of the crops are intimately related to the cycles in the activity of industry, and that it takes between one and two years for a good or bad crop to produce the maximum effect upon the activity of the pig-iron industry."[2]

"If the cycles of the yield per acre are correlated with the cycles

[1] On the relationship of Probable Error to r, see note 2 on p. 455 and the discussion on Probable Error, Chapter XI.

[2] Moore, H. L., *Economic Cycles: Their Law and Cause*, pp. 109–110.

of general prices, we find, for a lag of three years in general prices, $r = .786$; for a lag of four years, $r = .800$; for a lag of five years, $r = .710$. The cycles in the yields per acre of the crops are, therefore, intimately connected with the cycles of general prices, and the lag in the cycles of general prices is approximately four years." [1]

Professor Persons employs the Pearsonian coefficient of correlation in his recent study of a business barometer.[2] The purpose of the study is to construct a business barometer. Of its uses he says:

"Economists and sociologists need such a barometer when dealing with the phenomena of a dynamic society; government officials when handling the problem of unemployment or when considering the advisability of inaugurating large government undertakings; manufacturers and dealers when considering the desirability of making extensions to their plants or of contracting or expanding their purchases, sales, or commitments; bankers need a business barometer to guide them in extending or calling their loans and discounts; and investors need one to direct their purchases and sales of securities." [3]

By computing the coefficients of correlation between cycles of relative wholesale prices and various series of statistics

[1] *Ibid.*, p. 122.

[2] Persons, Warren M., "Construction of a Business Barometer Based upon Annual Data," *American Economic Review*, December, 1916, pp. 739–769.

[3] *Ibid.*, p. 739. The need for interannual correlation is indicated in a recent article by J. Arthur Harris. He says: "Practically such means of prediction as correlation and regression formulæ should find wide application in breeding operations where it is desirable to weed out or send to the butcher at the earliest possible moment those individuals which cannot be kept with the maximum profit. If the correlation between the egg production of a fowl in her pullet year and her laying capacity in any subsequent year be high, it is clear that those which on the average are to prove unprofitable may be sent to the pot when most desirable for that purpose, and before they have consumed two or more years' feed without yielding the maximum return in eggs. If, on the contrary, there be no correlation, the labor of selection in the pullet year is an unnecessary expense. If a cow's milking capacity be closely correlated with her milking record in her heifer year, the culling of dairy herds may be profitably carried out in the first year. In plant breeding experiments, involving either sexual or vegetative reproduction, selection of individuals for future propagation must be made, and at as early a date as possible. If the future yield per plant of hay can

indicating business conditions, when the price series precedes and lags behind the others, Professor Persons selects nine series as a business barometer. These with the coefficients for various periods are shown in the following table:

TABLE G

COEFFICIENTS OF CORRELATION BETWEEN CYCLES OF RELATIVE WHOLESALE PRICES AND CYCLES OF SERIES ENTERING INTO THE BUSINESS BAROMETER, 1879–1913 [1]

SERIES CORRELATED WITH RELATIVE WHOLESALE PRICES	COEFFICIENTS OF CORRELATION PRICES PRECEDE (–) OR LAG BEHIND (+) BY:						
	–2 yr.	–1 yr.	0 yr.	+1 yr.	+2 yr.	+3 yr.	+4 yr.
	+	+	+	+	+	+	+
Gross receipts of railroads	.847	.917	.945	.856	.748	.637	—
Net earnings of railroads	.690	.763	.862	.839	.803	.811	—
Coal produced	.787	.865	.931	.880	.795	.731	.630
Exports from the U. S. .	.547	.671	.783	.786	.772	.328	—
Imports into the U. S. .	.796	.796	.861	.754	.578	.445	—
Pig-iron produced . . .	—	—	.756	.738	.631	.617	.528
Price of pig-iron . . .	.406	.558	.763	.739	.637	.576	—
Immigration[2]	.606	.718	.789	.626	.494	—	—
Relative wholesale prices	.811	.923	1.000	.923	.811	.691	.548

By the same method he found other series, such as "shares sold on the New York Stock Exchange, new railroad mileage, the percentage of business failures," [3] in which the maximum

be estimated with considerable accuracy from a first year's culture the process of selecting clonal strains can be carried out with far greater rapidity than if one must wait for the results of subsequent years' tests. In all such cases the finality of a first judgment must depend in large degree upon the closeness of correlation between the results of successive experiments — in short upon the value of the inter-annual correlation coefficient." Harris, J. Arthur, "The Value of Inter-annual Correlations," in *The American Naturalist*, Vol. XLIX, November, 1915, p. 707.

[1] *Op. cit.*, p. 757.

[2] Fiscal year. Calendar year for all other series.

[3] *Op. cit.*, p. 765.

correlation occurred one year in advance of the business barometer, and which are, therefore, useful in forecasting business conditions. An extension of the same method permits the correlation of series for shorter periods and suggests the possibilities of calculating a sensitive business forecaster. It is hoped that Professor Persons' intention to make this more detailed calculation will be realized.

(2) Application of the Coefficient of Correlation to Frequency Series

The following examples show the application of the coefficient of correlation to frequency series. The first is worked out as in the historical series above; the second, in the form of a correlation table.

In an address on *Concentration of Power Supply,* Mr. Samuel Insull, President of the Commonwealth Edison Company, Chicago, said in relation to statistics there considered: "The income per kilowatt hour goes down pretty steadily, the output per capita goes up pretty steadily, the load factor improves as selling price is lowered, and the output per capita goes up as the selling price is lowered." [1] These conclusions were based upon a consideration of the United States Census figures for 1912 on the generation of electrical energy giving the capacity load factor,[2] output per capita, and income per kilowatt hour by states. It is the comparison of the first and the next to the last of these ratios — load factor and income per K.W.H. — which is tested out by the use of the correlation formula.[3]

[1] Address before the Finance Forum of the Young Men's Christian Association, New York, 1914, privately printed, p. 26.

[2] Ratio of average load to capacity in this case, p. 26.

[3] These figures are clearly inadequate for a satisfactory study of this character, but are used here simply as illustrative of the uses to which data

Following the plan used above in historical series, the following table gives the original facts, and the necessary computations for the coefficient of correlation:

TABLE H

TABLE SHOWING BY STATES THE CAPACITY LOAD FACTOR AND THE INCOME PER KILOWATT HOUR IN THE GENERATION OF ELECTRICAL ENERGY

State	Capacity Load Factor %	Deviations from Average Load Factor x	Deviations Squared x^2	Income per K.W.H. (in cents)	Deviations from Average Income per K.W.H. y	Deviations Squared y^2	Product of Deviations (x's) and (y's)
Total . .	av. 21.4		4144.61	av. 3.45		177.2011	− 444.735
Alabama . .	22.7	+ 1.3	1.69	2.49	− .96	.9216	− 1.248
Arizona . .	25.4	+ 4.0	16.00	3.56	− .11	.0121	− .440
Arkansas .	12.4	− 9.0	81.00	5.45	+ 2.00	4.0000	− 18.000
California .	33.9	+ 12.5	156.25	1.59	− 1.86	3.4596	− 23.250
Colorado . .	25.3	+ 3.9	15.21	2.89	− .56	.3136	− 2.184
Conn. . . .	19.2	− 2.2	4.84	4.10	+ .65	.4225	− 1.430
Florida . .	12.5	− 8.9	79.21	5.11	+ 1.66	2.7556	− 14.774
Georgia . .	17.8	− 3.6	12.96	2.01	− 1.44	2.0736	+ 5.184
Idaho . .	37.0	+ 15.6	243.36	1.37	− 2.08	4.3264	− 32.448
Illinois . .	29.3	+ 7.9	62.41	2.52	− .93	.8649	− 7.347
Indiana . .	19.9	− 1.5	2.25	3.26	− .19	.0361	+ .285
Iowa . . .	14.4	− 7.0	49.00	6.45	+ 3.00	9.0000	− 21.000
Kansas . .	22.0	+ .6	.36	2.19	− 1.26	1.5876	− .756
Kentucky .	15.9	− 5.5	30.25	3.64	+ .19	.0361	− 1.045
Louisiana .	10.9	− 10.5	110.25	12.25	+ 8.80	77.4400	− 92.400
Maine . .	22.7	+ 1.3	1.69	1.74	− 1.71	2.9241	− 2.223
Maryland .	5.0	− 16.4	268.96	1.37	− 2.08	4.3264	+ 34.112

may be put by those who desire to trace out similar relationships in business. If data existed for individual plants as units rather than for whole states, the correlation undoubtedly would be more marked.

TABLE H *Continued*

State	Capacity Load Factor %	Deviations from Average Load Factor x	Deviations Squared x^2	Income per K.W.H. (in cents)	Deviations from Average Income per K. W. H. y	Deviations Squared y^2	Products of Deviations (x's) and (y's)
Mass. . .	17.5	− 3.9	15.21	4.17	+ .72	.5184	− 2.808
Mich. . . .	23.2	+ 1.8	3.24	2.19	− 1.26	1.5876	− 2.268
Minn. . .	22.7	+ 1.3	1.69	3.72	+ .27	.0729	+ .351
Miss. . . .	14.6	− 6.8	46.24	4.02	+ .57	.3249	− 3.876
Missouri . .	21.7	+ .3	.09	4.18	+ .73	.5329	+ .219
Montana . .	58.0	+ 36.6	1339.56	1.05	− 2.40	5.7600	− 87.840
Nebraska .	18.6	− 2.8	7.84	4.98	+ 1.53	2.3409	− 4.284
Nevada . .	48.6	+ 27.2	739.84	1.38	− 2.07	4.2849	− 56.304
New Ham.	25.0	+ 3.6	12.96	1.84	− 1.61	2.5921	− 5.796
New Jersey	24.4	+ 3.0	9.00	2.85	− .60	.3600	− 1.800
New Mex.	12.9	− 8.5	72.25	5.50	+ 2.05	4.2025	− 17.425
New York .	32.1	+ 10.7	114.49	2.63	− .82	.6724	− 8.774
N. Car. . .	18.7	− 2.7	7.29	1.90	− 1.55	2.4025	+ 4.185
N. Dakota	12.9	− 8.5	72.25	7.01	+ 3.56	12.6736	− 30.260
Ohio . . .	18.6	− 2.8	7.84	2.99	− .56	.3136	+ 1.568
Oklahoma .	19.7	− 1.7	2.89	4.54	+ 1.09	1.1881	− 1.836
Oregon . .	20.7	− .7	.49	2.39	− 1.06	1.1236	+ .742
Penn. . . .	15.7	− 5.7	32.49	4.14	+ .69	.4761	− 3.933
Rhode Island	18.4	− 3.0	9.00	3.71	+ .26	.0676	− .780
S. Carolina	30.7	+ 9.3	86.49	1.24	− 2.21	4.8841	− 20.553
S. Dakota .	14.0	− 7.4	54.76	4.58	+ 1.13	1.2769	− 8.362
Tenn... . .	17.4	− 4.0	16.00	3.24	− .21	.0441	+ .840
Texas . . .	27.6	+ 6.2	38.44	3.38	− .07	.0049	− .434
Utah . . .	26.0	+ 4.6	21.16	1.75	− 1.70	2.8900	− 7.820
Vermont . .	21.9	+ .5	.25	2.07	− 1.38	1.9044	− .690
Virginia . .	8.1	− 13.3	176.89	2.65	− .80	.6400	+ 10.640
Wash. . .	14.2	− 7.2	51.84	4.33	+ .88	.7744	− 6.336
West Va. . .	16.1	− 5.3	28.09	2.60	− .85	.7225	+ 4.505
Wisconsin .	24.9	+ 3.5	12.25	2.92	− .53	.2809	− 1.855
Wyoming .	16.1	− 5.3	28.09	6.24	+ 2.79	7.7841	− 14.787

The standard deviation of the x series is 9.39, and of the y series 1.95. r by the formula, $\frac{\Sigma(xy)}{n\,\sigma_1\,\sigma_2}$, is $\frac{-444.735}{860.593}$, or $-$ 0.517, and the probable error, $\pm$ .0721. That is, correlation is negative and significant since it is approximately 7 times the probable error.[1] On the basis of the coefficient the generalization of Mr. Insull seems warranted. As to whether it is "an absolute demonstration of the necessity of monopoly in the production and distribution of energy"[2] is another question and one upon which no judgment is passed.

A convenient method of calculating the degree of correlation between two series is by means of what is known as a double-entry or double-frequency table. Examples of such tables are those illustrating the results of dice throwing given above.

"Each row in such a table gives the frequency-distribution of the first variable for cases in which the second variable lies within the limits stated on the left of the row. Similarly, every column gives the frequency-distribution of the second variable for cases in which the value of the first variable lies within the limits stated at the head of the column. As 'columns' and 'rows' are distinguished only by accidental circumstances of the one set running vertically and the other horizontally, and the difference has no statistical significance, the word *array* has been suggested as a convenient term to denote either a 'row' or a 'column.'"[3]

The manner in which the coefficient of correlation is determined for data arranged in this manner is indicated in the following double-frequency table — Table I — com-

[1] See the discussion of probable error, *supra*.
[2] *Op. cit.*, p. 27.
[3] Yule, G. Udny, *An Introduction to the Theory of Statistics*, pp. 157 and 164.

paring assessed values of improvements and of lands for 300 parcels of real estate in the city of New York. [1]

The question upon which an answer is desired is: In the sections chosen,[2] do relatively high or low improvement values go with relatively high or low (or the reverse) land values? The data are arranged as in the illustration of dice throws, each piece of property being placed in the table according to a double characteristic — assessed value of improvements and of land. No decided tendency is shown for the data to arrange themselves in a compact area extending from the upper left to the lower right or from the lower left to the upper right hand corners. That is, by inspection, neither marked positive nor negative correlation is present, the two characteristics being apparently independent, and the instances scattered about pretty much at random.

By applying the Pearsonian formula, r is found to equal $+ .0976$ and the probable error, $\pm .0002$. That is, the correlation is positive but negligible. The way in which r is calculated for such a series is as follows: Arithmetic means and standard deviations are determined in the usual manner. Columns indicated as d, d^2, and fd^2 in both series are used to find the standard deviations, the arithmetic means in this case being computed separately. In order, however, to calculate the products of the deviations from the arithmetic means — that is the (xy's) in the two series, it is necessary to treat the differences from their respective arithmetic means as made up of several parts rather than as single quantities. For instance, the item $+ 208.41$, in the column marked $\Sigma\,(xy)$, is obtained by multiplying each

[1] Data taken from Haig, Robert Murray, *Some Probable Effects of the Exemption of Improvements from Taxation in the City of New York.* New York, 1915, pp. 145–150.

[2] Upper East Side tenement and Rivington Street sections.

TABLE

TABLE CORRELATING ASSESSED VALUES OF IMPROVE-

Assessed Values in (000's) Dollars	Land 5–10	10–15	15–20	20–25	25–30	30–35	35–40	40–45	45–50	50–55	55–60	60–65	f Frequencies (Total)
Improvements 0– 5	—	4	10	1	—	—	1	—	—	—	—	—	16
5–10	11	8	19	30	5	1	—	—	1	—	—	—	75
10–15	13	9	23	38	6	1	—	—	—	—	—	—	90
15–20	11	—	—	14	2	—	—	—	—	—	—	—	27
20–25	5	—	4	1	2	4	1	—	—	—	—	—	17
25–30	20	17	2	1	1	—	—	3	1	—	—	—	45
30–35	2	9	3	—	1	—	—	3	2	—	1	1	22
35–40	—	—	—	—	1	1	—	1	—	—	—	—	3
40–45	—	—	—	—	2	—	—	—	—	2	—	—	4
45–50	—	—	—	—	—	—	—	—	—	—	—	—	0
50–55	—	—	—	—	—	—	1	—	—	—	—	—	1
Frequencies (Total) f	62	47	61	85	20	7	3	7	4	2	1	1	300
Deviations from Arith. Mean d	-11.28	-6.28	-1.28	$+3.72$	$+8.72$	$+13.72$	$+18.72$	$+23.72$	$+28.72$	$+33.72$	$+38.72$	$+43.72$	
Deviations Squared d^2	128	39	2	14	76	187	350	562	824	1136	1498	1910	
Deviations Squared Times Frequencies fd^2	7,936	1,833	122	1,190	1,520	1,309	1,050	3,934	3,296	2,272	1,498	1,910	27,870

Arith. Mean $= 18.78$ (Land)
S. D. or $\sigma_1 = 9.55$ (Land)

I

MENTS AND LAND — 300 PARCELS IN NEW YORK CITY

DEVIATIONS FROM ARITH. MEAN d	DEVIATIONS SQUARED d^2	DEVIATIONS SQUARED TIMES FREQUENCIES fd^2	PRODUCTS OF THE RESPECTIVE DEVIATIONS IN THE TWO SERIES $\Sigma\ (xy)$
− 13.73	188	3,008	+ 212.54
− 8.73	76	5,706	+ 8.73
− 3.73	14	1,260	+ 93.99
+ 1.27	2	54	− 69.29
+ 6.27	39	663	+ 208.41
+ 11.27	128	5,760	− 2,508.70
+ 16.27	262	5,764	+ 2,226.38
+ 21.27	454	1,362	+ 981.82
+ 26.27	686	2,744	+ 852.20
+ 36.27	1318	1,318	+ 678.97
		27,639	+ 2,685.05

Arith. Mean = 16.23 (Improvements)
S. D. or σ_2 = 9.60 (Improvements)

$$r = \frac{+\ 2685.05}{300 \times 9.55 \times 9.60} = +\ .0976$$

P. E. = ± .0002

of the items in the series 5, 4, 1, 2, 4, 1, by their corresponding differences from the arithmetic mean in one series — that is, (− 11.28), (−1.28), (+ 3.72), (+ 8.72), (+ 13.72), (+ 18.72) — and these in turn by + 6.27 — the difference which the total is from the arithmetic mean in the other series, thus:

$$\left.\begin{array}{l} 5 \times -\ 11.28 \\ 4 \times -\ \ \ 1.28 \\ 1 \times +\ \ \ 3.72 \\ 2 \times +\ \ \ 8.72 \\ 4 \times +\ 13.72 \\ 1 \times +\ 18.72 \end{array}\right\} \times +\ 6.27 = +\ 208.41$$

The other items in the product-difference column are similarly obtained. The total, + 2685.05, is the numerator for the correlation formula.

The grouping of data within a table serves as a sufficient index of the degree of correlation when it is marked. When data are widely scattered, it is necessary to use some graphic or numerical method of measuring it. A rather involved method is that followed in Table I. Less involved ones are in common use. For instance, Bowley, in correlating daily maxima and minima temperature changes by means of a double-frequency table, says:

> "If there is correlation, it will be found that the medians or arithmetic averages of each row form a regular progression, and similarly for each column." [1]

Mathematical and graphic means of measuring correlation are fully treated by Yule,[2] Elderton,[3] Bowley,[4] and others. It is not our intention to describe them here. Our purpose

[1] Bowley, A. L., *Elements of Statistics*, p. 323.
[2] Yule, G. U., *An Introduction to the Theory of Statistics.*
[3] Elderton, W. Palen, *Frequency Curves and Correlation.*
[4] Bowley, A. L., *Measurement of Groups and Series.*

is rather to illustrate in a simple way the meaning of correlation and to indicate its use in business and general economic fields.

The simple straightforward method employed above suffices for the construction of business barometers and forecasters, and for correlating trade, banking, and other phenomena with price movements. For more specialized uses, reference must be made to more detailed studies.

IV. Conclusion

Any comparison of phenomena having to do with economics and business is inherently difficult. The *other things* which so frequently are held to be equal in the natural sciences refuse to obey any well-defined law in matters relating to the social sciences. Comparison is particularly difficult when reliance is placed largely, if not solely, in statistics and statistical methods. Too frequently, the desire for statistical regularity and conformity is so dominant that the limitations of both statistics and statistical method are forgotten or ignored. It is inadequate simply to test the appropriateness of statistical devices. It is the condition back of these affecting the origin, methods of collection, tabulation, etc., which must be kept in mind. Units of measurements, coefficients, statistical abbreviations of all sorts, etc., must be scrutinized for errors, bias, non-application, due to change in time, place, and conditions. At every step it is necessary continually to bear in mind the statistical cautions which apply and to realize the limitations of the statistical approach.

Narrow cause-and-effect relations should not be expected. As has been shown, causes and effects are rarely exhibited singly. Any attempt to seek an absolute cause and an absolute effect is in a large degree futile. Most studies involve

correlation rather than narrow causation, and it is important that this truth be extended to fields of business and general economics. The problems associated with them are complex both as to cause and effect. The ways in which they are exhibited differ for time and for place. A realization of this on the part of the student or business man will prevent an undue optimism from characterizing the zeal with which he attempts to prove or disprove a complex thesis by faulty data and simple statistical means.

The statistical is one phase of the inductive method. In the analysis of problems, in the establishment of laws, it is a means and not an end. "Statistics," so called, is almost solely method. The great need to-day in business circles is an appreciation of the significance of facts and a familiarity with the ways in which they may be used to develop rules of guidance. Statistical facts, while complete in many fields, are far from satisfactory in others. But they are too often regarded simply as records of past performance, rather than as live, functioning indexes of future policy and possibilities. The outlook has been directed more to the past than to the future. But criticism applies not only to statistics, but equally as much to the use or lack of use which is made of them. To ignore a fact, statistical or otherwise, is never justified. A realization of this truth would go far toward putting statistical methods in the same favorable light as that now occupied by accounting. It is hoped that the volume here contributed will in some small degree help to accomplish this end.

References

Bowley, A. L. — *Elements of Statistics*, pp. 316–334.

——, *Measurement of Groups and Series*, pp. 61–74, "Correlation between Two Groups"; pp. 82–88, "Correlation between Series."

Davenport, E. — *Principles of Breeding*, Ch. 13, pp. 453–472.
Elderton, W. P. — *Frequency Curves and Correlation*, Ch. VI, pp. 106–125.
Elderton, W. P. and E. M. — *Primer of Statistics*, Ch. 5, pp. 55–72.
Kincer, J. B. — "A Correlation of Weather Conditions and Production of Cotton in Texas." *The Monthly Weather Review*, February, 1915. Vol. 43, pp. 61–65. (U. S. Dept. of Agriculture, Weather Bureau.)
King, W. I. — *Elements of Statistical Method*, Ch. XVI, pp. 186–197; Ch. XVII, pp. 197–216.
Moore, H. L. — *Economic Cycles: Their Law and Cause*, Ch. V.
Pearson, Karl. — *The Grammar of Science*, Chs. IV and V.
Persons, W. M. — "Correlation of Economic Statistics," *Publication of the American Statistical Association*, Vol. 12, 1910, pp. 287–322.
——. "The Construction of a Business Barometer Based upon Annual Data," in *American Economic Review*, December, 1916, pp. 739–769.
Whipple, Guy Montrose. — *Manual of Mental and Physical Tests*, Ch. III, pp. 14–40, particularly.
Yule, G. U. — *Introduction to the Theory of Statistics*, Ch. IX, pp. 157–190; Ch. X, pp. 191–206.

SUBJECT INDEX

(The citations are to pages and include only the subject matter. A *Personal* index is appended separately.)

PERSONAL INDEX

(The numbers refer to pages and include citations in both text treatment and references.)

Printed in the United States of America.